Economic Order Quantity (EOQ) Inventory Management - Essays in Experimental Economics

WEI, SIJIA

Economic Order Quantity (EOQ) Inventory Management - Essays in Experimental Economics

Sijia Wei

Department of Economics and Finance

Durham University

This thesis is submitted for the degree of

Doctor of Philosophy

St Mary's College

2019

Declaration

I, Sijia Wei, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work. No part of this thesis has previously been submitted else where for any other degree.

Chapter 2, Chapter 3 and Chapter 4 of this thesis are based on joint research papers with Professor Jason Shachat and Dr. Jinrui Pan - 'Cognitive stress and learning economic order quantity inventory management: An experimental investigation' and 'Cognitive reflection and economic order quantity inventory management: An experimental investigation'. All three authors contributed significantly to this project.

05/12/2019

Abstract

This thesis consists of six chapters to experimentally study aspects of how levels of individuals' cognitive stress, cognitive ability and self-regulatory resource affect their decision making under the Economics Order Quantity (EOQ) inventory management environment.

In Chapter 3 we use laboratory experiments to evaluate the effects of cognitive stress on inventory management decisions in a finite horizon economic order quantity model. We manipulate two sources of cognitive stress. First, we vary participants' participation in a pin memorization task. This exogenously increases cognitive load. Second, we introduce an intervention to reduce cognitive stress by only allowing participants to order when inventory is depleted. This intervention restricts the policy choice set. Increases in cognitive load negatively impact earnings with and without the intervention, with these impacts occurring in the first year. Participants' in all treatments tend to adopt near optimal policies. However, only in the intervention and low cognitive load treatment do the majority of choices reach the optimal policy. Our results suggest that higher levels of multitasking lead to lower initial performance when taking on new product lines, and that the benefits of providing support and task simplicity are greatest when the task is first assigned.

In Chapter 4 we use laboratory experiments to evaluate the effects of individuals' cognitive abilities on their behaviour in a finite horizon economic order quantity model. Participants' abilities to balance intuitive judgement with cognitive deliberations are measured by the Cognitive Reflection Test (CRT). Participants then complete a sequence of five "annual" inventory management tasks with monthly ordering decisions. Our results show that participants with higher CRT scores on average earn greater profit and choose more effective inventory management policies. However these gaps are transitory as participants with lower CRT scores exhibit faster learning. We also find a significant gender effect on CRT scores. This suggests hiring practices incorporating CRT type of instruments can lead to an unjustified bias.

In Chapter 5 we use laboratory experiments to evaluate the effects of individuals' ability to self-regulate on inventory management decisions in a finite horizon economic order quantity model. An ego depletion task is implemented aiming to diminish one's self-regulatory resources. From a psychological point of view, self-control is impaired when the mental resource has been used up over effortful control of responses. In our experiment, participants complete an ego depletion task followed by a sequence of five "annual" inventory management tasks with monthly ordering decisions. Our results show there is no obvious treatment effect on participants' self-regulation ability.

Acknowledgements

The journey of pursing a doctoral degree has been the most inspiring gift life has given me. I have a long list of people to acknowledge who have helped me along this journey, it would not have been so fruitful and pleasant without their love and company.

I am extremely grateful to my supervisor Professor Jason Shachat for his invaluable guidance and continued support during my PhD studies. Throughout the years Jason has provided me with a lot of professional suggestions for improvement, encouragement for my ideas, challenges and inspiration to keep me motivated. I would also like to thank my second supervisor Dr. Jinrui Pan for her patience, encouragement and expertise. Thanks for their genuine belief in me to develop our research from scratch to a high standard of completion.

Besides my supervisors, I would like to thank my annual reviewers Dr. Kenju Kamei, Dr. Riccardo Mogre, Dr. Min Liu and my thesis committee for taking time to read my work and giving insightful comments on the materials in this thesis.

We gratefully acknowledge the financial support from Durham University Business School. Many thanks to the help and support from the Experimental and Behavioural Economics Lab within Newcastle University. I also greatly appreciate the support from Mr. Hangkun Wu and Ms. Xin Su for taking part in our interviews with corporations in the field.

I am very thankful to my PhD colleagues Dr. Yanjun Tan, Dr. Zhixian Qu, Dr. Qinjuan Wan, Mr. Matthew Walker, Ms. Hongming Zhao, Dr. Handing Sun and Mr. Xiao Liang at Durham University for their help with coding, testing and running the experiments. And of course, for their good friendship that has made the past three years such an enjoyable time.

I will always appreciate the help and encouragement from my former university advisors Professor Miguel Fonseca, Professor Dieter Balkenborg, Professor Rajiv Sarin, Dr. Cecilia Chen, Dr. James Watson and Dr. Susan Long in the past few years.

I would like to express my gratitude to my teaching team colleagues Dr. Simone Tonin, Dr. Frankie Chau and Dr. Ahmed Elsayed, and my lovely students at Durham University, especially IFE class 2017-18. You have made the days we spent together part of my very best memories in Durham.

My sincere thanks also goes to my dear friends, Ms. Yilin Shi, Ms. Mengyan Zhao, Ms. Xiaorui Qian for being my family over the past ten years when I have lived so far away from my parents, and Dr. Yiwen Cui, Ms. Yanjun Liu, Ms. Qiao Ma, Ms. Luchang Yong, Ms. Yinan Zhao, Ms. Shuhan Liu, Ms. Bingran Li, Ms. Yuanqi Lin, thank you for sharing every moment of my life, my joys and gripes.

With all my heart, a very deep appreciation goes to my parents who supported me both financially and mentally throughout these years, my grandma for sharing with me wisdom from the Bible, and also my beloved late grandpa, I am extremely fortunate to have you in my life. Thank you for your love, understanding, and always encouraging me to explore the fascinating world of behavioural economics.

Contents

List of Figures

List of Tables

1 Introduction and Literature Review

Inventory management is the key to the success of supply chain operations. It refers to the process of controlling and supervising the ordering and storage of a company's inventory. The objective of inventory management is to maintain the optimum number of the inventory in storage, and to provide uninterrupted services at a minimum cost. An efficient handling of the inventory process is critical to the growth and stability of business operations.

The newsvendor model and the economic order quantity (EOQ) model are the two fundamental models used in supply chain management to determine the optimal inventory levels. A large number of existing literature analysed decision biases in inventory management in the newsvendor setting (Bolton & Katok, 2008; Bostian, Holt, & Smith, 2008; Feng, Keller, & Zheng, 2011). In a typical newsvendor's experiment, subjects are facing static price and uncertain demand, and order decisions are made repeatedly for multiple rounds. The good is perishable in the newsvendor problem, and will lose its value by the end of the decision rounds. Other important inventory management models have also drawn the attention from extensive empirical experiments, for example, the multi-level supply chain beer game for durable good with certain demand (Croson, Donohue, Katok, & Sterman, 2014; Narayanan & Moritz, 2015; Niranjan, Wagner, & Bode, 2011), and the (S, s) inventory management model for durable good with uncertain demand (Khaw, Stevens, & Woodford, 2017; Magnani, Gorry, & Oprea, 2016). Results from existing studies show that even under the circumstances where there are no distractions, individuals systematically make suboptimal inventory management decisions.

In the present research, we will approach inventory managers' ordering behaviours from a less complex model. The experiments are designed based on the EOQ model where the good is not perishable, and demand for a product is constant across decision periods. The EOQ model eliminates the effect of risk averse, demand chasing and strategic concerns on subjects' behaviours. Subjects can make decisions on how many units to order in each decision period. There is a holding cost for any inventory held in storage and a fixed ordering cost every time an order is placed. The EOQ model provides a formula to work out the replenish point which minimises the sum of the inventory holding costs and ordering cost; and the ordering amount which avoids stockouts or excess inventory in storage. The EOQ is a prevalent tool of inventory managers in the field. We have conducted interviews with inventory managers from several Chinese durable goods manufacturers, which confirmed the widespread adoption of the EOQ built-in Economic Resource Planning (ERP) systems as their predominant tools for inventory decision support. Later in the thesis we will provide evidence on the managerial value of our research study.

In Chapter 3, we introduce a finite horizon deterministic EOQ environment as our basic experimental framework. Our baseline treatment is called the "Unrestricted", where participants can order additional inventory each month regardless of the current inventory level. Around the baseline, we adopted a two-factor experimental design. First, we examined exogenous shocks to participants' cognitive load by adding a concurrent task that competes for the inventory manager's cognitive resources; second, we examined the effect of restricting participants to only ordering once inventories are depleted, which removes the possibility of violating the optimal inventory policy. We observe there is a trend that participants learned to adopt near optimal

EOQ policies in general. Our results show that experimental participants perform less optimally when there is a competing task or when the intervention is absent. However, these negative impacts occur predominantly when participants first face the inventory decision problem.

Chapter 4 focuses on individual heterogeneity in terms of cognitive reflection, which is measured by the Cognitive Reflection Test (CRT). The CRT is a widely used performance-based measure designed to assess individual's tendency to override impulsive but wrong responses, and engage in a more effortful and reflective thinking for the right answer. We followed the same EOQ experimental settings with a different set of parameter values. We find that participants with higher CRT scores tend to make more effective inventory management decisions. However, the performance gap closes with experiences, as participants with lower CRT scores exhibit faster learning rate in realizing the expected profits. Consistent with previous literature, male participants have higher level of cognitive ability comparing to female participants. Our results suggest that the common human resource procedure of using CRT as criteria to filter job applicants would lead to gender biased selections. This practice discriminates against female applicants, when in fact their performance will not differ from the males on average after building initial experiences from training activities.

In Chapter 5, we examine self-control and the impact of ego depletion upon decision quality in inventory management. Self-control is a limited amount of mental resource that is exhaustible over effortful control of responses. A person would be considered to be at a state of ego depletion that would affect his self-control abilities on subsequent tasks. Previous research studies show that under stochastic inventory management processes subjects will often adopt the optimal inventory policy but then their choices subsequently move away. Our experimental design provides a more static inventory management environment where the deterministic demand allows for consistent feedback. The treatment variable we implement is the level of ego depletion subjects experienced in a letter 'e' task. We anticipated that the ego depletion task matters ex-ante. However, we do not observe self-control issue in our static experimental setting.

The rest of Chapter 1 will provide a comprehensive review on relative literature.

1.1 Inventory management models

Researchers observe behaviour biases in dynamic supply chain management environments that contain strategic considerations such as the beer game. Niranjan et al. (2011) pointed out that there are limitations in people's behaviours when responding to delayed responses. The supply line underweighting (SLU) theory describes that individuals' overreactions, in this case, the over ordering behaviour has been primarily attributed to the fact that inventories in-transit are not easily recognised. The authors tested the SLU theory in an experimental setting and suggested that it is not the only behavioural bias that is influencing the ordering behaviour. Cognitive limitations and bounded rationality are driving biased responses because they were not explicitly explained to people. SLU falls in the category when individuals are incapable of making correct decisions, whereas correction behaviour applies when communication is not allowed or the buyer has limited power. Individuals tend to exhibit over ordering behaviour to influence upstream suppliers when they are less powerful players in the supply chain system.

Croson et al. (2014) conducted an experimental research on the bullwhip effect that offered an explanation based on coordination risk. Coordination risk occurs when the decision rules for each supply chain managers are not known to each other with certainty. Firstly, in order to test purely on the behavioural causes of the bullwhip effect, the experimental design eliminates possible computing errors arise from the operational causes of the bullwhip effect. The results suggest that decision errors occur because participants were uncertain about the behaviour of other supply chain partners. Therefore, participants deviate from the equilibrium to hedge the risk. Significant improvements were made on the participants' performance in the treatment group where coordination risk was eliminated. Lastly, coordination stock was introduced to prevent managers from coordination failures, which has also alleviated the effect.

Bostian et al. (2008) replicated the newsvendor "pull-to-centre effect" in laboratory experiments. The authors state that the intuition behind this effect is that the bias tends to pull the order quantity towards the amount of the recent demand, and subjects are more likely to focus on the results of the order decisions they made other than the results of forgone decisions. The design of the treatments varies in terms of frequency of feedback and order decisions. Consistent with Schweitzer and Cachon (2000), orders are generally below the optimum order level under high-profit margin settings, while orders are generally above the optimum order level under low-profit margin settings. Strong pull-to-centre effect was found in both high frequency and low frequency decision designs. The pull-to-centre effect also exists in designs with doubled payoffs, which indicates that the deviation is not related to payoff sensitivity. It has been proved by previous studies that the deviations from optimum order quantity cannot be explained by risk aversion or the prospect theory. The authors showed in their research that the learning model which concerns both heuristics and bounded rationality explains the order behaviours better.

Under the newsvendor model, subjects face the risk of overage cost by ordering too many, while also face the risk of shortage cost by ordering too few. Wang and Webster (2009) proved in a single-period newsvendor problem that loss-aversion leads to decision biases. Risk-averse subjects' marginal utility of underage loss and overage loss are different, they are more sensitive to losses than gains. As a result, risk-averse subjects tend to order more than risk-neutral subjects. In the research by de Vericourt, Jain, Bearden, and Filipowicz (2013) the authors observed gender differences in participants' ordering behaviours. Male subjects tend to order more in high margin settings, and the gender difference is partially driven by subjects' risk attitudes. Males are in general more risk taking than females. Long and Nasiry (2014) further proved the fact that prospect theory can be used in explaining the newsvendor problem.

Rudi and Drake (2014) investigated several behaviour issues in newsvendor problem, namely, level behaviour, adjustment behaviour and observation bias. Level behaviour refers to subjects' average order level; adjustment behaviour refers to the degree of adjustment from period to period. Subjects participated in a laboratory experiment where level of good margin and the degree of feedback were manipulated as treatments. In order to study decision makers' observation bias, in the uncensored demand group, subjects observe actual quantities sold from the feedback; while in the censored demand group, subjects observe excess orders and stockouts, but only given information about actual demand quantities when overage occurs. Consistent to previous studies, the pull-to-centre effect is observed when demands are not censored. Behaviours observed under censored demand violate the pull-to-centre effect and lead to lower order quantities in general. The authors conclude that adjustment behaviour causes more profit erosion than level behaviour, and full demand information given in feedback can lead to worse performance in making inventory management decisions.

When operating in a stochastic environment, such as inventory dynamics, agents are expected to adjust continually in order to re-optimising. (S, s) model shows that when it is costly to make changes, agents should make adjustments following the "state dependent" rule. In state dependent models, the best strategy is to ignore small volatilities and only adjusts to re-optimise when the delays become significantly unprofitable. In an experimental study by Magnani et al. (2016), subjects failed to perform under the "state dependent" rule under dynamic settings, instead, subjects exhibit "time dependent" adjustment behaviours. The researchers argue that participants' bounded rationality and the substantial cognitive costs from responding to the dynamic environment are the key drivers of such behaviour. In another recent experimental research, Khaw et al. (2017) studied about discrete adjustment behaviour in laboratory where subjects make decisions to maximise reward in a dynamic environment. Contrary to the prediction under the (S, s) model, the adjustments didn't follow the constant optimal policy even when the adjustment is free at any time. A model of rational inattentive adjustment can better explain the observed pattern. They stated that the discrete adjustment behaviour implies participants' inattention, cognitive limitations and their willingness to track the changing environmental state. This finding may well be contributing to the understanding of discrete adjustment behaviours in other economical settings.

Stangl and Thonemann (2017) studied how different performance metrics from supply chain management that include equivalent information can affect the performance of actual human inventory decision makings. Days of supply is the length of time period that inventory is held in stock, while inventory turn rate is the inverse that measures the frequency at which a company replaces its inventory annually. One of their experiments is based on the basic EOQ model that analysed the cost allocation behaviour between order placing and inventory holding. In classic

EOQ model, decision variable for inventory managers to decide is normally the order quantity. The authors in this study asked subjects to determine the ordering cost for products with different holding costs. Therefore, the higher the ordering cost they chose, the more frequent the orders are placed, the lower corresponding holding cost would be. Subjects received an endowment before they made their decisions, and the total costs were then deducted from the endowment. The results indicate that for all products, subjects are willing to spend more on ordering costs under the inventory turn rate metric than under the days of supply metric, with the optimum cost optimization solution lies in the middle. Since the cost function is steeper below than above the optimal solution, the average total costs under the days of supply metric are significantly higher than those under the inventory turn rate metric.

A more recent study based on EOQ model by K.-Y. Chen and Wu (2017) is designed to examine inventory management learning behaviour under both stable and changing environment. The experiment consists of two parts, in Phase I participants make ordering decisions under fixed cost parameters repeatedly, and in Phase II participants make ordering decisions with changing operational costs. There are three treatments subject to the type of cost that changes across rounds, each participants participated in one of the three treatments. In treatment "VarK", keeping the holding cost stable, ordering cost is altered in each period; in treatment "VarH", holding cost is altered instead when ordering cost remains stable; and in treatment "VarKH", both holding cost and ordering cost are changing. Participants can view their historical information and results while making ordering decisions. With deterministic feedbacks, the inferior choices from past decisions can be observed with certainty. The authors conclude that learning occurs over periods, and subjects learn much faster about the optimal choice under stable environment than under changing environment. In contrast to evidence from newsvendor's experiments where subjects tend to repeat suboptimal decisions, suboptimal decisions tend not to be repeated with deterministic feedbacks. A new learning model is proposed to explain this behaviour as individuals can learn from past experiences and reduce evaluation errors of potential decisions over time; in addition, learning is influenced by randomness occurred in the environment.

1.2 Cognitive stress

1.2.1 Choice complexity

Bode and Wagner (2015) raised the question 'Which supply chain characteristics result in most of the disruptions in supply chain management'. Disruptions are usually unanticipated events that stop the normal stream of inventory moving along the supply chain, consequently affect the operation and financial stability of the firm. Two forms of complexities associated with supply chain disruptions are terms as structural complexity and dynamic complexity. The former term refers to the condition when there are multiple items in the supply chain system, and the later describes the interactions between the items. The degree of supply chain complexity has been increased unsurprisingly in recent years. It has been attributed to the growing demand of skills and knowledge of supply chain mangers, the increasing hierarchical level within organizations and the geographically spreading of the logistic operating points. The drivers of complexity often act interdependently, which in turn, amplifies the level of stress on inventory managers and the frequency of disruptions in supply chain operations.

Many choices were made without evaluating the whole set of available options in everyday decision making, consequently, best options may be missed. Caplin, Dean, and Martin (2011) explained using a satisficing model how the decision quality was influenced by incomplete consideration of the choice set. The experiment task requires participants to select monetary rewards displayed as addition and subtraction operations, and involves treatments varied in terms of number of choices within the choice set and the level of complexity of these items. The research shows that participants often fail to find the best options, and they tend to stop searching once the reservation level has been realised. Participants are observed to make increasing mistakes when facing larger and more complicated choice sets. The observed search orders typically follows the sequential search theory, whereas levels of complexity also affect search orders. Subjects in general exhibit top-to-bottom searching mode, and there are also simple-to-complex searchers.

The standard revealed preference theory applies when decision makers consider the full set of feasible options. The theory is limited in its explanatory power if only a subset of the options were considered. In real world scenarios, decision makers exhibit limited attention and select the most preferred option within the bounded consideration set. Masatlioglu, Nakajima, and Ozbay (2012) studied limited attention on revealed preference. The bounded consideration set is referred as attention filter. The research provides a framework to distinguish between preference from attention and inattention, and provides policy implications to real world markets. For example, consumers have cognitive limitations when facing the whole range of products on the market. Products compete with each other for consumers' limited attention. Under the attention filter assumption, the unpopularity of a product in the market can be attribute to either it is not well evaluated by consumers, or it hasn't attracted consumers' attention.

Choice complexity affects people's behaviour and result in increasingly suboptimal economical decision making. Abeler and Jäger (2015) analysed in a controlled laboratory experiment, how the complexity tax system influence the reaction to additional tax rules. In the experiment, subjects are paid with piece rate and are paying income taxes. The complexity of the tax system differs in terms of the number of distinct tax rules subject faces. The simple system is designed to serve as a baseline where participants understand the tax rules and the incentives. The complex system includes more tax rules that characterised the feature of real tax systems. The results show that less optimal decisions were made under the more complex experimental setting; participants underreact to additional tax rules and did not make sufficient adjustments towards the new profit-maximising strategy. Further analyses also show that underperformance was largely driven by participants with lower cognitive abilities.

In the research study by Lleras, Masatlioglu, Nakajima, and Ozbay (2017), the authors used a choice model to analyse the 'more is less' effect where an excess amount of available options can result in less desirable decision making performance. Consumers tend to have restricted attention focused on only a subset of the choice set and fail to consider other alternatives in the set. Contrary to the prediction under classical economic models where more choices is always better, too much choice can result in choice overload. Consumers tend to ignore more options when the choice set becomes larger, consequently, make decisions from an even smaller portion of the whole choice set. The increasing number of available options overwhelms people and exposes them to cognitive load. Under such cognitive limitations, consumers choose their reduced subset following different types of heuristics which can be welfare-reducing.

Bolton and Katok (2008) examined the effect of experience and feedback on newsvendor problem in three experimental studies. In the first study, subjects were given extra decision making rounds and smaller sets of ordering options. The results suggested that experience is more effective in enhancing performance than giving participants statistical feedbacks. Personal experience does a better job in correcting pre-conceived biases. Subjects with learning experience could achieve 90% of the optimum profits. However, only little improvement was achieved when the option sets are thinner. In the second study, subjects who can observe the profits given by the foregone options did not improve in their overall performance. In the third study, subjects who were restricted to make standing orders performed better in terms of expected payoffs and reaching optimum ordering point. In order words, restricting decision makers from responding to short-term fluctuations improves their inventory management performance.

In the cross-national newsvendor's study by Feng et al. (2011), the authors revisited the 'thinning the set of options' treatment in Bolton and Katok (2008). They argue that the extremeness aversion can explain the reason why reducing the number of options in the choice set didn't improve participants' decision making performance in the previous research. In the current design, more options were added in to the choice set, making the optimal options no longer the extreme values in the choice set. The results from the experiment show that thinning the set of order options in a way that the optimal order quantity is not an extreme option in the choice set does lead to better performance. The results also show that Chinese subjects behave differently from American subjects, as they exhibit more explicit pull-to-centre effect.

Tokar, Aloysius, and Waller (2012) studied the debiasing effect in a series of supply chain inventory replenishment experimental tasks. The first experiment includes single-echelon replenishment tasks with debiasing treatment components such as knowledge of bullwhip, inventory position, and use of a target order-up-to quantity. In the second experiment, they introduced multiple supply lines to increase the complexity of the coordination risk. Participants were first presented with an explanation of the beer game mechanism, and were given written instructions and a sheet to keep their decisions in record. Their results suggest that additional quantity of information from the multi-echelon setting has increased participants' cognitive load. Participants were observed to have difficulties processing the written instructions due to cognitive limitation. It has been proved in the research that behavioural biases can result in inventory replenishment inefficiencies. In this case, the available information is more than the sufficient amount of information that can help participants to make effective decisions.

1.2.2 Concurrent task

In psychology, multitasking is referred to as a 'task-switching' between multiple operations from time to time, and there is usually a cost linked to changing between ongoing tasks. Buser and Peter (2012) designed an experiment to study how people's productivity was affected when working on multiple tasks. The authors argue that the designs in previous studies were not ideal to investigate the questions because the tasks were either too simple, or not incentivised, or subjects were not allowed to switch at anytime. In their experiment, participants were assigned into three different treatment groups where they work on puzzles in a fixed time period. Group one participants were asked to complete the tasks in a determined order, group two were required to multitask, and group three can switching freely between tasks. Participants in the second group performed worse than the first group, however, group three also performed badly. Therefore, the authors conclude that scheduling is an effective way to increase productivity. And the result didn't confirm the popular stereotype that women are better at multitasking.

Miller (1956) first introduced the method of using the PIN number task to exogenously impose cognitive load. It was referred to as people's capacity to transmit information. The experiment was designed to measure the amount of input information and the amount of the transmitted information when the input has been increased. The amount of transmitted information is expected to increase with the amount of input information at the beginning and gradually become less accurate after the critical point. The purpose of the research is to find this critical point as the upper limit of the human response to the stimuli. Experimental results show that discontinuity happens at the number of seven which indicates the limit on our attention span. Subjects can subitize of any categories below seven, and were said to estimate above seven. The magical number of seven was observed and used extensively in psychology experiments as the limit on the capacity of human memory.

In a food choice experiment by Shiv and Fedorikhin (1999), a number pin task was used to constrain on participants' processing resources to test the affect and cognition in consumer decision making. Participants were presented with a choice decision making between chocolate cake and fruit salad. The chocolate cake was a more superior choice on affective dimension but less favourable on the cognitive dimension while the fruit salad is the opposite situation. The experiment adopted between-subject design where the high processing resource group were asked to memorise a two-digit number and recall later, and the low processing resource group were given a seven-digit number. As a result, choices were influenced by the level of constrain imposed on participants' processing resources. Chocolate cake was a more popular choice when participants were making decisions under high level of constrains. Participants with sufficient processing resources were better at avoiding more affective reactions.

Roch, Lane, Samuelson, Allison, and Dent (2000) studied using a two-stage model the impact of cognitive load on subjects' overconsumption behaviour of scarce resource. The model suggests that subjects would initially engage in an automatic process where they anchor their choices following the 'equal-division' heuristic. Further, subjects with sufficient cognitive resources would engage in a more systematic process to adjust their decisions to a self-serving direction. In the experiment, participants were assigned into groups and could request recourses from the group's recourse pool. Participants under high cognitive load were given an 8-digit number to remember. The experimental results supported the model prediction that participants under low cognitive load requested more from the recourse pool, and the amount requested by low cognitive load participants has greater variability. Participants under high cognitive load were lack of cognitive resources in the second stage to process the self-serving amount that they would like to request in addition to the equality amount they have chosen in stage one.

S. Allred, Duffy, and Smith (2016) designed a series of laboratory experiment to test the assertion that cognitive load increases the central tendency bias. It has been widely accepted that cognitive resource is bounded, thus, the manipulation of cognitive load is predicted to reduce subjects' available working memory and diminish the accuracy in decision making tasks. The authors used an adjustment task where subjects adjust the length of a line displayed on the screen to match with the line they saw earlier. In the mean time, subjects were engaged in a cognitive load tasks that require them to remember either a 2-digit code or a 6-digit code before the adjustment task and recall the code afterwards. The codes were either numbers or letters under different experiment settings, and the results proved that both numbers and letters were effective in inducing cognitive load. The authors also found that cognitive load results in central tendency bias, and higher cognitive load exacerbates such bias.

S. R. Allred, Crawford, Duffy, and Smith (2016) further analysed the relationship between cognitive load and strategic behaviours. The authors performed a within-subject analysis of cognitive load manipulation to study individual behaviours under different levels of cognitive load. The within-subject design also eliminated the effect caused by different payments across treatments. In the experiment, subjects played different forms of one-shot strategic games while memorizing a number code. In the low load condition, the number code was a 3-digit binary number, and in the high load condition, the number code was a 9-digit binary number. The results suggest two opposing effects: in consistent with previous studies, subjects under high cognitive load exhibit a reduced ability to reach the optimal decision; whereas they also tend to perform more sophisticated since they believed that they have a cognitive ability disadvantage comparing to their partners. Thus, higher level of cognitive load may not necessarily lead to less sophisticated choices.

Duffy and Smith (2014) studied strategic behaviour in the finitely repeated multi-player prisoner's dilemma game. The authors induced a differential cognitive load in subjects by requiring them to memorize different lengths of number stings. In each period, subjects in the low cognitive load treatment were told to remember a 2-digit number; in the high cognitive load treatment, subjects were told to remember a 7-digit number. Then after making the choice in the game, they were asked to recall the number. The authors found that the manipulation of cognitive load has a negative affect on subjects' cognitive ability. Subjects under the low cognitive load condition performed more strategic defection choices near the end of the game, and were better able to condition their behaviour on outcomes from previous history.

Hinson, Jameson, and Whitney (2003) proposed that people's available working memory condition affects their decision making on intertemporal choices. The temporally myopic behaviour is defined as acting in favour of short-term rewards over long-term rewards. In the first experiment, the authors measured how subjects value immediate reward versus delayed. In order to impose different levels of cognitive load upon subjects' working memory, one group of the participants were asked to keep a string of five numbers in memory, while another group of the participants engaged in a random number selection task. The authors observed from the experimental result that maintaining a number string in memory has reduced the perceived value of delayed rewards, whereas, the impact was not significant with the random number generation group. In the second experiment, subjects were presented with additional options of the delayed periods. The results show that additional options also increased participants' cognitive load, thus, reduced the perceived value of delayed rewards. Further experiment using real monetary rewards shows that the behaviour of delayed discounting is more significant than hypothetical rewards.

Different methods were implemented in experimental studies to test how cognitive load would affect logic, math ability, and strategic behaviour. Neys (2006) used a classic dot memory task in their experiment. Subjects were solving syllogistic reasoning tasks while their executive resources were burdened by memorizing a dot pattern. Under the low-load condition, subjects were asked to remember a simple three dots pattern that filled in a 3x3 matrix, and recall the pattern later in the experiment. A more complex four dots pattern is filled in the 3x3 matrix under the high-load condition. In the syllogistic reasoning task, the believability of the conclusions were either consistent with or in conflict with the logic. The results reveal that the executive burden under the high-load condition is heavier comparing to the burden under the low-load condition. Further, the memorizing dual task reduced syllogistic reasoning task performance when beliefs were in conflict with logic.

The research by Sprenger et al. (2011) examined the effect of cognitive load on probability judgment. In their experiment, subjects were required to remember and later recall a list of 1-8 letters while completing a probability judgment task at the same time. The judgement task asked subjects to make probability judgment using a 0 to 100% scale on the probability of a given item on the menu would be ordered by a given customer. The results indicate that less accurate judgement decisions were made under cognitive load, and the magnitude of judgement increases with the magnitude of cognitive load.

Rydval (2012) studied how does cognitive load interfere with economic behaviours. The author introduced a time-series forecasting task in a laboratory experiment where participants make forecasts based on observations on seasonal patterns and forecasting errors from the decision history. The level of cognitive load was manipulated by requiring subjects to memorize the relevant data related to the forecasting task. The between-subjects experimental design includes a low memory load treatment and a high memory load treatment. Consistent with the prediction, the average performance on the forecasting task is better in the less memory-intensive treatment.

Benjamin, Brown, and Shapiro (2013) conducted a laboratory experimental study using high school students to examine the impact of cognitive load on math ability. Participants were required to remember a 7-digit number while solving SAT level math quizzes, and recall the number string after the quiz. Both correct answers in the math quiz and correctly recall the number string were incentivized with monetary payoffs. The results reveal that the performance on the math quizzes was negatively affected by cognitive load. However, the reduction on performance was not statistically significant, which implies a weak manipulation of cognitive load.

Carpenter, Graham, and Wolf (2013) used Hit games and standard beauty contest game in their research to identify the relationship between cognitive ability and strategic sophistication. The authors observed the influence of cognitive load on strategic behaviours in one of their laboratory experiments. Subjects were asked to play Hit games and standard beauty contest game while simultaneously engaging in memorizing a random 7-digit number string. Raven's matrices were used to measure subjects' cognitive ability. In general, cognitive load caused less sophisticated behaviours in games. Moreover, in the Hit games, the performance of higher ability player reduced even significantly than the reduction in the performance of lower ability players.

There are plenty of experimental studies examined the effect of cognitive load under various situations. Deck and Jahedi (2015) summarized in their study the findings from previous literature which impose exogenous cognitive load directly on economic tasks. They also conducted new experiments where subjects were making decisions on a variety range of economic tasks while remembering either a 1 or 8-digit code. The results suggest that cognitive load reduces subjects' numerical accuracy and increases the likelihood of anchor. Moreover, the authors argue that the subjects who are more sensitive to the manipulation of cognitive load drive this effect the most.

1.3 Cognitive ability

The CRT test was first introduced by Frederick (2005) to measure a type of cognitive ability. The test consists of three questions where each of the questions has an intuitive but wrong answer. The questions are easy to understand once explained, however, reaching the correct answers requires subjects to exert more effort to resist the impulsive wrong answers. Research evidence shows that the intuitive answers are dominating other wrong answers given; and the intuitive answers were often considered first even by the participants who gave the correct answers. The average score among college students participated in the experiment was 1.24 out of 3. Higher CRT scores are proven to be associated with more patience in intertemporal choices, and people with lower CRT scores are more prone to risk taking behaviours in the loss domain than in the gain domain. In general, men score higher in the CRT test then women. Interestingly, the observed correlation between CRT scores and time preferences is more significant for women, and the correlation with risk preferences is more significant for men.

Oechssler, Roider, and Schmitz (2009) investigated in an experimental study if CRT test is related to economical and financial behavioural biases, in particular, conjunction fallacy, anchoring and conservatism. The experiment was conducted over the Internet. Participants were given a set of behavioural questions mixed with CRT questions. Participants who gave 0 or 1 correct answer in the CRT test belong to the 'low' group, as they are more impulsive decision makers; the 'high' group consists of more reflective decision makers who answered 2 or 3 questions correctly. The results show that individuals in the 'low' group are more likely to exhibit biased behaviours. However, anchoring was observed among all participants, not in particular more significantly related to participants with lower CRT scores. Hoppe and Kusterer (2011) further examined behavioural biases that are potentially related to CRT, including susceptibility to the base rate fallacy, conservatism bias, overconfidence, and the endowment effect. Similarly, participants who gave 0 or 1 correct answer are in the 'low' group, participants who gave 2 or 3 correct answers are in the 'high' group. Their findings show that subjects with lower CRT scores are more likely to exhibit both behavioural biases to overweight or underweight the base rate. In the overconfidence task, participants were asked to estimate the number of correct answers they have given in a general knowledge quiz. Participants in the 'low' group were observed to be overconfident in the estimation. However, the endowment effect did not differ in terms of CRT scores.

In a more recently experimental study by Cueva et al. (2016) studied how CRT test score correlates with behavioural choices in a wide variety of tasks. The subjects were classified into three groups according to their performance on the CRT task. Subjects who gave at least two correct answers are in the 'reflective' group, subjects who gave at least two intuitive answers are in the 'impulsive' group, and the rest are in the 'residual' group. Physiologically, they have observed that the CRT score is negatively correlated with the 2D:4D ratio; psychologically, neuroticism and extraversion are significantly negatively correlated with the CRT score. Subjects in the 'impulsive' group are more likely to be inequality averse, and they are less likely to act consistently in the lottery choice task. The experimental results also show that female participants gave significantly less correct answers than male, and they tend to give impulsive answers more frequently when they make mistakes. Further analysis shows that the overall gender difference was largely driven by females. The positive correlation between subjects' performance on the CRT test and their financial literary no longer exists when control for the gender parameter.

Bergman, Ellingsen, Johannesson, and Svensson (2010) studied cognitive ability and anchoring in the context of consumer behaviours. Participants were asked to express their maximum willingness to pay decisions on six different products. The elicitation process follows the BDM procedure. In the experiment, both psychometric test and CRT test were used to test for individuals' cognitive abilities. They have found that anchoring – cognitive ability relationship exists, anchoring decreases when individuals' have higher cognitive ability scores. However, the CRT test measures a weaker correlation than psychometric test. In the research by Brañas-Garza, Garcia-Muñoz, and González (2012), the CRT test was proven to be a better predictor than the Raven test in the beauty contest game. Subjects were asked to play six rounds of the beauty contest game, where different multipliers were used in each round, followed by the Raven test and the CRT test. Their results show that subjects with higher level of cognitive abilities exhibit higher level of reasoning, and tend to play more dominant strategies in the beauty contest game.

Brañas-Garza, Kujal, and Lenkei (2015) discussed more CRT related literature in a meta-survey study. The research aims to study whether implementation methods affect individuals' performance on the CRT test. They conclude that participants' performances did not differ with or without monetary reward; less CRT questions were answered correctly if there were preceding tasks taken place, which proves that cognitive resources decline over time and effort. Overall, student subjects outperformed non-students, while male subjects outperformed females. Finally, subjects performed marginally better in computer based experiments.

CRT test has been proved to be a good predictor for subjects' performance on heuristics-and-biases tasks Toplak, West, and Stanovich (2011). The experiment includes Vocabulary and Matrix Reasoning subtests to measure cognitive ability; 15 heuristics-and-biases tasks to test for rational thought; two sets of syllogistic reasoning tasks for biased beliefs; set shifting, inhibition and working memory test to measure executive functioning; and a self-report questionnaire for thinking depositions. The results suggest that CRT test result is correlated with both cognitive ability and rational-thinking skill. Regression analysis shows that CRT test predicts rational-thinking skill not purely attribute to intelligence measures, but also executive functioning and thinking dispositions. CRT test provides a direct performance measure to assess the tendency towards reasoning errors from miserly processing rather than intelligence tests or self-reported measures.

Liberali, Reyna, Furlan, Stein, and Pardo (2012) answered in their research paper that CRT is not just another numeracy scale. The study measures an individual's objective numeracy scales by testing basic probability questions and mathematical concepts; subjective numeracy scales by an eight-item self-perceived numerical competency; a three-item CRT test for an individual's cognitive reflection; a short form of the Raven Advanced Progressive Matrices test for cognitive capacity; and an individual's ability of inhibitory control using a 'go/no-go' task. The results support that CRT captures one's ability of monitoring, editing and active responding. The distinct feature of the CRT test is that wrong answers also come across smarter people's mind, however, they inhibit and edit the answers. The bivariate correlations between CRT and general numeracy suggest that they are not simply equivalent to each other.

Campitelli and Gerrans (2014) used a mathematical approach to study whether cognitive reflection test is actually measuring cognitive reflection. They have raised the concerns that the CRT test may only measure mathematical ability since it consists of three math problems. The mathematical model was designed to capture the hierarchical structure of the CRT test. The model includes both inhibition parameter and mathematical parameter, where the inhibition parameter measures the ability of resisting the impulsive and intuitive response, and the mathematical parameter measures the ability of following adequate mathematical processes. The results show that CRT test is not equivalent to a pure mathematical test, it also measures one's ability of rational thinking and the tendency of involving in an actively open-minded thinking. The approach also enabled the authors to identify gender differences in the CRT test. The inhibiting behaviour of male subjects is related to both rational thinking and open-minded thinking abilities, whereas only rational thinking is linked to the behaviour of female subjects.

In psychology research, people often use the number of correct responses from the CRT test to assess individual's level of cognitive reflectiveness. Pennycook, Cheyne, Koehler, and Fugelsang (2016) discussed the validity of the opposite strategy that some researchers choose to use the number of incorrect responses from the CRT test as a measure of intuitiveness. The authors analysed how did the two cognitive traits, namely, reflectiveness and intuitiveness associated with subjects' self-report measures of intuitive–analytic cognitive styles. Subjects were asked to rate their 'need for cognition' and 'faith in intuition' on a 1-5 scale. The results show that the participants' cognitive styles can be predicted by their CRT-reflective measures. However, the observed correlation between the numbers of CRT intuitive answers and the participants' self-report measure of intuitiveness was not robust. It was also not clearly explained in previous literature how intuitiveness affects people's performance in the CRT test. Therefore, CRT test is valid at measuring one's reflective ability, but not intuitive thinking.

Primi, Morsanyi, Chiesi, Donati, and Hamilton (2016) argued that the popularity of the CRT test might be limited in practice among less educated people due to its difficulty. A large number of participants scored 0 in previous experimental studies. Their research looks into the psychometric properties of the CRT test, aiming to develop a new version of the test that can provide a more accurate scale for all education backgrounds. Firstly, the authors used the item response theory analyses to examine the psychometric properties of the CRT test, and found that the test has a highly discriminate nature over subjects with different levels of cognitive ability. A longer 6-item scale was designed for more precise measurement on both higher level and lower levels of cognitive ability. The newer version of the CRT test results is correlated with participants' intelligence, decision-making and reasoning skills, thinking dispositions and numeracy abilities at a similar degree as comparing to the original CRT test. Consistent with previous studies, gender differences were observed from the newer version test results, and the differences were more significant among younger age.

Psychometric tests were designed to measure primarily on candidates' cognitive ability, its growing popularity has been widely accepted by recruiters in many industries. Schmidt and Hunter (1998) summarized a large selection of theoretical and practical research findings on predicting job and training performance. The top three combinations of methods used were general mental ability test plus work sample test, integer test, and structured interview. The level of cognitive ability is recognised as a good predictor for both entry-level applicants and experienced hires. Economically, the efficiency of selecting on the basis of candidates' cognitive ability has been directly reflected in the proportional increase of the final output.

Harper (2008) discussed the strengths and weaknesses of psychometric tests and suggested that it may be wrong for companies to blindly focus on the test results and stereotype employees' behaviour. Such expectation raised the awareness of both recruiters and employees. Also, there have been existing arguments over the consistency of the cognitive ability test results over time. Earlier work noted the potential trade-off between workplace diversity and performance. Companies have explicitly used cognitive ability tests in their recruiting process as it provides a good prediction on levels of productivity. However, often result in adverse outcomes in workplace diversity. Newman and Lyon (2009) demonstrated the advantages of selecting on the basis of both cognitive ability and conscientiousness in order to meet the workplace diversity goal while maintaining the overall productivity. In the case where companies may value these two factors with different weight, the recruiting strategy can be adjusted accordingly to satisfy both goals simultaneously. This type of recruiting strategy has also been proven to increase the percentage of hires and average productivity from the underrepresented group.

There are existing studies prove that participants' CRT score explains their quality of inventory decision-making. Supply chain performances are shown to depend upon an individual's level of cognitive reflection when managing multi-level supply chain in a beer game (Narayanan & Moritz, 2015). The research study suggested that the CRT profile of decision makers explains the bullwhip effect generating upstream the supply chain. For all Level-2 thinkers a much better job is done in terms of costs, order variances and demand amplification. Level-2 thinking can intervene to override the immediate Level-1 type of intuitive answers, in particular for inexperienced decision-makers. Participants with higher cognitive reflection tend to anticipate their incoming supply line and all converge to a more efficient decision while participants with lower cognitive reflection tend to ignore the interactions. Standard mitigation strategies are proved to have intervened the performance of teams with high cognitive reflection participants.

The performance also differs based on individual's level of reflective thinking when managing perishable goods in newsvendor's experiment (B. B. Moritz, Hill, & Donohue, 2013). The experimental subjects involve both supply chain professionals and business school students. They have found that individuals with higher cognitive reflection exhibit a lower tendency to chase demand and that the cognitive reflection is a better predictor of performance outcomes. In the condition where there is a variation in supply chain demand, individuals with high cognitive reflection scores are observed to outperform those with low cognitive reflection scores in forecasting the future demand (B. Moritz, Siemsen, & Kremer, 2014). Furthermore, decisions that were made either too far or too slow - correspond to under or over thinking - tend to result in bad performance. More reflective thinkers tend to exhibit quicker, and as well a more constant decision speed in giving the response. Therefore, the authors suggest that encouraging the participants to make the decisions in a moderate amount of time would improve the overall performance of supply chain forecasting.

1.4 Ego depletion

Baumeister, Bratslavsky, Muraven, and Tice (1998) conducted a series of experiments on ego depletion, aiming to answer the question if self-regulation is a limited resource. Ego depletion is defined as the depleting of one's willingness to engage in volitional actions due to the exertion of their self-control capacity. Activities, such as controlled processing, active choice, initiating behaviour, and overriding responses all require self-regulation resource. In the first experiment, participants who resisted the temptation from eating chocolates were less persistent in the subsequent task. In the second experiment, self-regulation was weakened when participants were required to take the responsibility for their own decisions on a counter-attitudinal speech. In the third experiment, suppressing emotions were also proved to have impaired the self-regulation resource. In the fourth experiment, the authors introduced a letter 'e' task. The ego-depletion group needs to cross out the letter 'e's following a specific rule where they will have to exert effort cognitively to override existing habits. The results suggest that participants were more passive if they have exhausted their self-regulation recourse in the preceding letter 'e' task.

Ego depletion has been proved to influence individuals' charitable behaviour (Fennis, Janssen, & Vohs, 2008). The authors developed a two-stage model to explain the social influence techniques behind the charitable requests. In stage one, the initial requests were presented to the consumers, which generate the foot-in-the-door (FITD) effect. Responding to these cognitively demanding questions consciously depletes individuals' self-regulatory resource, thus, weakened their active self. Therefore, the impaired self-control ability fosters the compliance-promoting heuristics (i.e., reciprocity, liking, and consistency). In stage two, individuals tend to follow heuristics to answer the requests due to the lack of self-regulatory resources, consequently, result in charitable giving behaviours. The ego-depletion tasks used in the experiments including the FITD technique, the Stroop task, the letter 'e' task, the geometric figure-tracing task, and the lowball technique.

Tyler (2008) studied whether monitoring for the relational value cues in a social environment consumes or depletes people's self-regulatory resources. Relational value cues indicate how much others value us as their social partners. In the self-regulatory depletion task, subjects were given unsolvable anagram puzzles to work on. They were expected to exert self-regulatory resources continuously whilst rearranging the letters until they find the correct combination to build a word. Since the tasks were designed to be unsolvable, individuals tend to give up when experiencing continuously failures. It requires self-control effort to override the impulse to stop in order to be persistent on the task. The experimental results show that relational value cues deplete self-regulatory resource. In the following experiment, subjects were given depletion tasks (i.e., the letter 'e' task, and the white bear task) prior the relational value monitoring activities. The results show that subjects were less effective at monitoring relational value cues, and were less able to identify more complex relations after the depletion task.

Wheeler, Briñol, and Hermann (2007) studied ego depletion and how it influences individuals' resistance to persuasion. Participants in the ego depletion condition were asked to complete the letter 'e' task. After the depletion task, subjects were presented with counter-attitudinal arguments. The depleted subjects were observed to give significantly more positive responses to the counter-attitudinal policy than the non-depleted group. They have concluded that resistance to persuasion diminishes with limited self-regulatory resource, especially when the arguments were made in weak statement. DeWall, Baumeister, Stillman, and Gailliot (2007) showed in their experimental research that lack of self-regulatory resources leads to more aggressive actions when individuals are insulted. The ego-depletion manipulation implemented in the research including food temptation, attention control, the Stroop task, the letter 'e' task, and to supress expressions. The results reject the argument that ego depletion increased aggressive impulses, and support that ego depletion limited individuals' inner restrains against impulsive aggressions. Muraven (2008) has also proved using the letter 'e' task that individual's level of self-control strength affects their control over the unintentionally stereotype acts. More depleted individuals tend to exhibit biased behaviours; in this case, they express greater prejudice attitudes. Participants who would not normally control their stereotypes intentionally were not affected by the ego-depletion treatment.

Hagger, Wood, Stiff, and Chatzisarantis (2010) discussed more ego depletion and self-control related literature in a meta-analysis research paper. The research provides analyses on whether different experimental conditions and implementation methods of the ego depletion tasks affect the effect size. They conclude that factors such as effort, perceived difficulty, negative affect, subjective fatigue, and blood glucose levels significantly affected the effect size. Further, depleting task duration, task presented by the same or different experimenters, whether or not reporting interim period, dependent task complexity, anticipation of future depletion task, adopting choice volition and cognitively demanding tasks could be used as moderator of the depletion effect. In addition, the effect size of Stroop types of depletion task is significantly smaller than the letter 'e' task.

2 Model

2.1 EOQ Model

To determine the optimum order quantity in inventory management, the traditional and most commonly used model is called the EOQ model. According to Harris (1913), economic order quantity (EOQ) is the optimum order quantity that minimizes the sum of inventory holding costs and inventory ordering costs. There are several assumptions underlying the simulation model used in the study:

1. Demand is determined at a constant rate, no safety stock is required;
2. Lead time is equal to 0, orders are received and delivered instantaneously;
3. Backorder is not allowed, any unsatisfied demand in a decision period will be considered a loss;

Under the EOQ model, the optimum order quantity Q is to minimize the total of inventory holding costs and inventory ordering costs. Ordering cost is a fixed cost for each order placed and does not depend upon the size of the order; inventory holding cost is also known as storage cost, it applies to the average quantity held in stock for a period of time. The basic cost function is:

$$TC = \frac{D * S}{Q} + \frac{h * Q}{2} \tag{1}$$

This is the definition of a few terms:

- TC = total cost
- D = demand
- Q = order quantity
- S = fixed cost per order
- h = monthly holding cost per unit

As shown in Figure 1, there is a trade-off between inventory holding costs and inventory ordering costs, resulting in a convex total cost curve. The optimum order quantity Q is the corresponding quantity where the holding cost curve intersects with the ordering cost curve, as well as where the slope of total cost curve is equal to zero.

Figure 2 describes the theoretical inventory order system under the EOQ model. An order of the economic order quantity Q is received immediately after an order has been placed and Q is sold at a constant demand rate. A new order will only be placed when the inventory level drops to the reorder point. The reorder point, in this case, is equal to 0 when the lead time is 0. Orders arrive instantaneously therefore there will not be any demand shortages. This inventory order cycle continues with the determined order quantity Q.

Figure 1: Classic EOQ Model

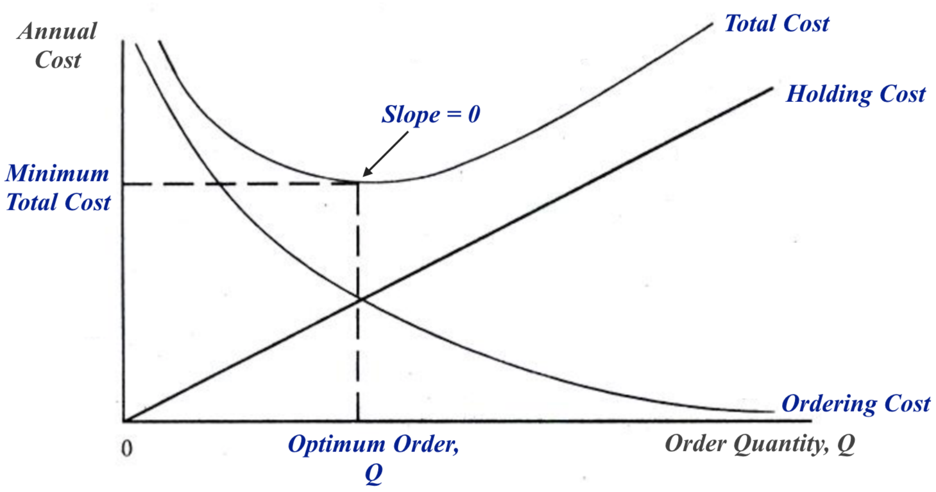
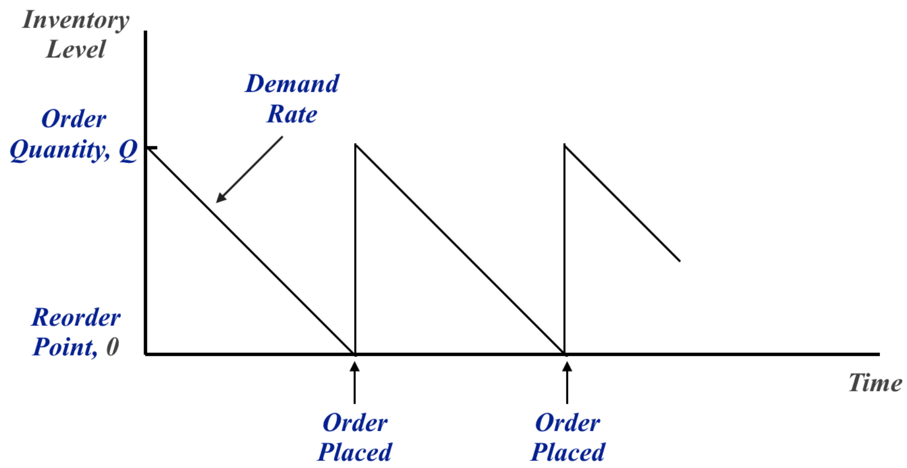


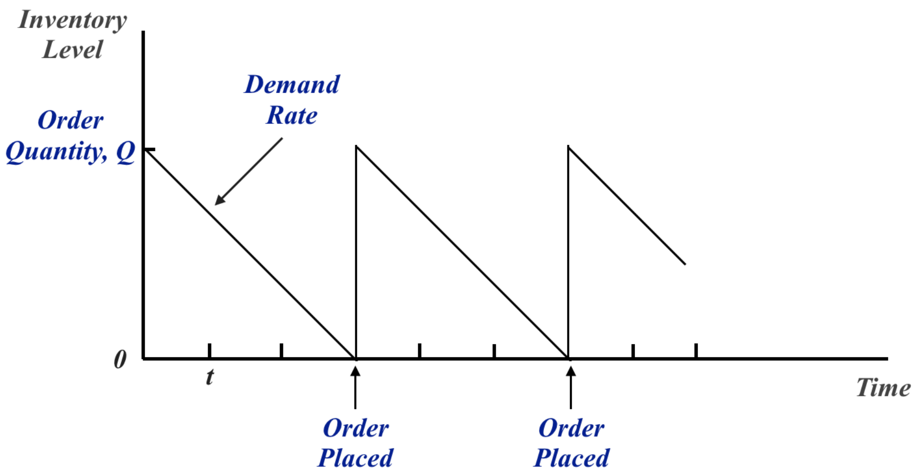
Figure 2: Economic Order Quantity Inventory Order Cycle



Consider one of the most commonly used replenishment systems, the periodic review system. Periodic review indicates that inventory status can be tracked on a regular periodic basis, and order decisions can be made at the beginning of each period. Figure 3 describes the periodic review inventory order system under the EOQ model. When the number of decision periods is fixed, time is equally distributed between reorder points. Orders of the economics order quantity Q are placed at each reorder point to bring the inventory level up to the optimum, with the ordering frequency equal to the optimum solution under EOQ model.

In this study, two order systems under the periodic review system were developed. One is the Unrestricted order system, orders of any amount can be placed at the beginning of each decision period; the other one is the Zero Only order system, orders can only be placed at the beginning of the decision periods where inventory levels equal to 0. The next section will discuss the fixed solution to the finite horizon EOQ model in our experiment that matches that of the infinite horizon.

Figure 3: Periodic Review System Inventory Order Cycle



2.2 Finite horizon EOQ Model

In the core decision-making part of our experiment, participants complete a series of six discrete dynamic inventory management tasks. We refer to each task as a year, indexed zero to five, and each year consists of twelve months, indexed by t. We use the following context to describe these tasks to a participant. The following set of parameter values in the EOQ model where used in Chapter 3.

The participant manages the enterprise 'S-store' which sells coffee makers with a constant demand rate (D) of 10 units per month. S-store sells a new model of coffee maker every year. Coffee maker orders are placed prior to the start of a month, an integer amount denoted q_t, and arrive without lag, hence are included in the calculation of a month's opening inventory. The participant chooses the quantity of each monthly order.

Monthly orders and demand determine the changing inventory levels. Let I_t denote the closing inventory for month t. The initial inventory of coffee makers prior to month one is zero, so the first month's opening inventory is the amount of the first month's coffee maker order, i.e. $I_0 + q_1 = q_1$. In general, the opening inventory of coffee makers in month t is $I_{t-1} + q_t$. This inventory is drawn down by the monthly sales, the lesser of the monthly order flow of 10 or the opening inventory (i.e. a stockout). This results in the closing inventory of $I_t = I_{t-1} + q_t - \min\{10, I_{t-1} + q_t\}$. When the model life cycle concludes at the end of month 12, any remaining inventory is disposed at no cost but also generates no revenue. Further, we limit a participant's monthly order by its annual demand, i.e., $q_t \in \{0, 1, 2, \ldots, 120\}$.

A participant's compensation, excluding a fixed show-up fee, is proportional to S-store's profits, which are expressed - as are all further monetary quantities - in experiment currency units (denoted ₱). Each coffee maker sells at a price of ₱7. So revenue in month t is $7 \cdot \min\{10, I_{t-1} + q_t\}$. S-store's cost has two components: a fixed ordering cost, S, of ₱45 whenever she places a strictly positive order; and a constant per-unit monthly inventory holding cost. The monthly inventory holding costs is calculated by multiplying the average inventory of coffee makers held in t, specifically $\frac{(I_{t-1}+q_t+I_t)}{2}$, and the monthly holding cost, h, of ₱1 per unit. The monthly profit of S-store is the difference between the revenue and costs, and is calculated

$$\pi_t(q_t, I_{t-1}) = \begin{cases} 7 \cdot 10 - S \cdot \mathbb{1}_{q_t>0} - \frac{I_{t-1}+q_t+I_t}{2} \cdot 1 & \text{if } I_{t-1} + q_t \geq 10 \\ 7 \cdot (I_{t-1} + q_t) - S \cdot \mathbb{1}_{q_t>0} - \frac{I_{t-1}+q_t}{2} \cdot 1 & \text{if } I_{t-1} + q_t < 10 \end{cases}$$

where, $\mathbb{1}$ is the indicator function.

A participant i's inventory *policy* for year a is the sequence of the twelve monthly quantity orders, $Q_{i,a} = (q_{i,1}, q_{i,2}, \ldots, q_{i,12})$. For a given inventory policy S-store's annual profits are,

$$\Pi_{i,a}(Q_{i,a}) = \sum_{t=1}^{12} \pi_t.$$

In the supply chain literature, the set of EOQ policies is the subset of inventory policies which only place a quantity order once inventory reaches zero with no stockouts allowed. In our dynamic decision-making environment, stockouts can occur if a non-optimal policy was chosen previously. Correspondingly, we adjust the definition of an EOQ policy to classify choices at these points off the optimal path.

(1). A participant only orders when the closing inventory of the previous period is less than 10 units, i.e., $q_t > 0$ when $I_{t-1} < 10$;

(2). A participant doesn't order when the closing inventory of the previous period is at least 10 units, i.e., $q_t = 0$ when $I_{t-1} \geq 10$;

(3). Participant's order guarantees no stockouts in t, i.e., $I_{t-1} + q_t \geq 10$.

The original EOQ model solution is derived assuming an infinite demand horizon, in which the average cost minimizing EOQ policy is to order the following quantity whenever the closing inventory of the previous period is zero,

$$q^* = \sqrt{\frac{2DS}{h}}. \tag{2}$$

In our context, the cost minimizing policy would be to order 30 coffee makers, an EOQ cycle length of three months, whenever closing inventory of the previous period is zero. This would also be the profit maximizing policy as average revenue is constant, up to the monthly demand capacity, and is greater than the minimum average cost. In our finite horizon setting the optimal policy does not change. However, if an inventory manager deviates from this policy early in the year, the optimal course can involve alternative EOQ actions later in the year.

Schwarz (1972) characterizes the optimal EOQ policies for the finite horizon of T months. First, we note the result that average total cost minimizing policy is to order according to Equation 2 if T is an integer multiple of the $\frac{q^*}{D}$. As simply following the EOQ policy of ordering 10 units each period is profitable in our environment, profit maximization will call for satisfying the full annual demand. The EOQ policy of always taking the EOQ action of 30 when inventory is depleted maximizes profit in addition to minimizing average cost.

As individuals do fail to act optimally, we now consider alternative decision horizons (i.e. shorter in this case). Let $C(T)$ be total incremental cost over the finite time interval T. We restrict our attention to policies which only place orders when inventory is zero. An EOQ cycle length is the interval of months between such orders, denoted by s_k, which is the interval between the $(k-1)th$ and the kth order. Let $C(s_k)$ be the total incremental cost for an EOQ cycle, and n be the number of orders over T. We can formulate the problem as

$$\min C(T) = \sum_{k=1}^{n} C(s_k) \qquad s.t. \quad \sum_{k=1}^{n} s_k = T$$

where

$$C(s_k) = S + hDt^2/2.$$

From the quadratic formulation, it is clear that in the optimal solution all of the s_k are of the same length. An EOQ constant inventory policy, denoted $\bar{Q}^{s_k}$, is one with a constant cycle length.

Let $C^n(T)$ be the total incremental cost for the interval T given n orders,

$$C^n(T) = nS + hDT^2/2n.$$

Minimising $C^n(T)$ gives

$$n^* = \sqrt{\frac{hDT^2}{2S}}. \tag{3}$$

Notice for the first month in our task, i.e. $T = 12$, this yields the same solution as the infinite horizon formulation, $n^* = 4$ and $s_k^* = 3$. Thus, the optimal order quantity $Q^* = 30$. Further investigations on situations when the horizon T is sufficiently small reveals that the optimal number of orders, n^*, is the smallest integer satisfying $n(n+1) \geq \dfrac{hDT^2}{2S}$.

An EOQ constant policy is an EOQ policy that orders the same quantity when an order is placed, and D/Q is an integer. In our experiment, the optimal policy is EOQ constant 30, $Q^* = 30$.

2.3 Dynamic solution to the finite horizon EOQ Model

Under our experimental design, there is a dynamic solution to the finite horizon EOQ Model conditional upon the starting month. Further investigations on situations when the horizon T is sufficiently small reveals that the optimal number of orders, n^*, is the smallest integer satisfying $n(n+1) \geq \dfrac{hDT^2}{2S}$. With the parameter values in Chapter 3, Table 1 gives an overview of the optimal solutions for different values of T.

Table 1: Optimal solutions for different T in our task

Month	T	$\frac{hDT^2}{2S}$	$n^*(n^*+1)$	The optimal order number (n^*)	The optimal EOQ cycle length (s_k^*) sequence	The optimal order size (q_k^*)
12	1	0.111	2	1	{1}	{10}
11	2	0.444	2	1	{2}	{20}
10	3	1	2	1	{3}	{30}
9	4	1.778	2	1	{4}	{40}
8	5	2.778	6	2	{3, 2}	{30, 20}
7	6	4	6	2	{3, 3}	{30, 30}
6	7	5.444	6	2	{3, 4}	{30, 40}
5	8	7.111	12	3	{3, 3, 2}	{30, 30, 20}
4	9	9	12	3	{3, 3, 3}	{30, 30, 30}
3	10	11.111	12	3	{3, 3, 4}	{30, 30, 40}
2	11	13.444	20	4	{3, 3, 3, 2}	{30, 30, 30, 20}
1	12	16	20	4	{3, 3, 3, 3}	{30, 30, 30, 30}

where, s_k^* is the optimal EOQ cycle length, and q_k^* is the optimal order size.

As we can conclude from the table, only in month 9 and 11, EOQ constant 40 and 20 become the optimal solution to the EOQ problem; EOQ constant 10 yields the most profits in month 12.

3 Cognitive Stress

3.1 Introduction

Inventory managers are responsible for monitoring and reporting on the inventory level of businesses. An efficient handling of the inventory process is critical to the success of business operations. This requires inventory mangers to possess great attention to detail and with excellent abilities to solve for dynamic optimization problems. Current workplace trends impose increasing demands upon these managers' cognitive resources (Ruderman, Clerkin, & Deal, 2017). Some examples of these trends are increasing complexity of supply chains (Bode & Wagner, 2015), the widely accepted increasing rates and scale of workplace distractions, and inter-department communications. Best inventory management practices require inventory managers to resolve complex sets of alternative solutions, and use their short-term memory to hold and process information about the past, present, and future values of key variables. We assess how increasing level of cognitive stress through the introduction of a concurrent competing task impacts decision-making quality, and how a strategic intervention can mitigate the impact of this stress.

An extensive literature shows that, even under the circumstances where there are no distractions, individuals systematically make suboptimal inventory management decisions. Decision-making biases and strategic considerations are often key factors diminishing individual performances in these tasks (Niranjan et al., 2011). When managing the inventory of a perishable good with uncertain demand, i.e. the newsvendor problem, decision makers neither follow the optimal risk neutral nor averse policies consistently in experimental studies. Bostian et al. (2008) observed the newsvendor "pull-to-centre effect" under different experimental settings in a laboratory study. When managing a multi-level supply chain for a non-perishable good with certain demand, participants generate large bullwhip effects in beer game experiments. Research has shown key factors driving the excessive inventory levels and variance include strategic uncertainty regarding other decision makers (Croson et al., 2014), limited level two thinking (Narayanan & Moritz, 2015) and failure to fully take account of the future deliveries of past orders. When managing a durable good with uncertain demand, optimal inventory management follows the (S, s) policy. Recent experimental studies (Khaw et al., 2017; Magnani et al., 2016) demonstrate that individuals take time to find the optimal policy, their policy adaptations are time dependent and often decisions deviate from the optimal once found.

In addition to all above mentioned inventory management environments, our research provides an extension on inventory managers' behaviour under the EOQ environment. Despite being one of the most commonly used models in operations management, behavioural studies have mostly overlooked it. The EOQ environment has several favourable features for our research question: participants have a relatively good chance of finding the optimal policy after repetitions; the solution is invariant to a decision maker's risk attitude; and, it is an individual decision problem absent of strategic considerations. Our finite horizon deterministic EOQ environment potentially possesses these properties. And we believe that the EOQ settings provide the best condition to begin an evaluation of how cognitive stress diminishes decision-making quality. We choose the parameters of our environment such that the optimal inventory policy of the finite horizon matches that of the infinite horizon; when inventory is depleted, the manager orders an optimal quantity that is the multiple of the monthly demand for the good (Schwarz, 1972). We refer to this multiple as an EOQ cycle length.

In most behavioural supply chain studies, participants do not determine 'when' to act. The EOQ solution in our environment is dynamic, as the manager doesn't make the same decision at each point in time. This gives us an opportunity to observe pure learning behaviour in a dynamic problem. Our baseline treatment is called the "Unrestricted", where participants can order additional inventory each month regardless of the current inventory level. Around the baseline we implement a 2 x 2 experimental design. Our intervention, the "Zero Only" treatment restricted participants from ordering when there is a positive level of inventory, which removes the possibility of violating the optimal inventory policy. The other factor we consider is the presence of a concurrent task that competes for the inventory manager's cognitive resources - we call this "High" treatment. There is no concurrent task in the "Low" treatment.

There is an a priori belief that our intervention will yield economically significant improvements. A growing and recent literature in economics, e.g., Abeler and Jäger (2015); Caplin et al. (2011); Lleras et al. (2017); Masatlioglu et al. (2012), examines and measures how individual choices are increasingly suboptimal as their choice sets increase in complexity. Bolton and Katok (2008) find that reducing the number of order options does not necessarily result in better performance for newsvendor decisions. However, Feng et al. (2011) observe that thinning the set of order options in a way that the optimal order quantity is not an extreme option in the choice set does lead to better performance. Our Unrestricted treatment corresponds to the case of an unsupported inventory manager, while the simplified choice set of the Zero Only treatment corresponds to active management intervention. This allows our experiment to provide evidence on the value of the EOQ practice.

The second factor we investigate is the presence of a concurrent task that competes for the inventory manager's cognitive resources. Tokar et al. (2012) finds experimental evidence of cognitive overload with an increased quantity of information. This would involve the introduction of inventory management responsibilities of additional product lines in the context of inventory management. However, managing the inventory of an additional product line typically introduces cross-demand impacts and potential synergies for inventory costs reductions. To control for the "costs" and "benefits" of successful inventory management we introduce an additional task unrelated to the inventory management one.

This concurrent task is the memorization of a PIN code at the beginning of each inventory year, and successful recall at the end of the year earns a monetary reward. The PIN task was first introduced by Miller (1956), and has been successively used in economics and psychology to exogenously shock cognitive load. Some recent examples of its application are in food choice (Shiv & Fedorikhin, 1999), generosity (Roch et al., 2000), strategic games (S. Allred et al., 2016; Duffy & Smith, 2014) and intertemporal choice (Hinson et al., 2003). Deck and Jahedi (2015) surveys the use of PIN task in economic experiments with financial incentives as well as reporting new experiments, one of which finds increasing PIN length reduces individual numeracy. To the best of our knowledge, we are the first to use this technique in behavioural operations management. Correspondingly, this allows our experiment to evaluate the impact of asking inventory managers to multitask.

Our results show that experimental participants earn more when the intervention is present or when there is no competing task. We observe there is a trend that participants learned to adopt near optimal EOQ policies in general. The restriction of managers to only place orders when inventories are exhausted and the alleviation of the competing task improved the chance for decision makers to reach the optimal inventory policy. It should be noted that these performance differences and suboptimal choices largely occur in the first three iterations of our environment. As performance improves rapidly across treatments, we attempt to characterize the individual-level learning driving this trend. We present the learning process as a decision tree which permits hierarchies of sophistication. We define the propensity to follow the basic characteristics of EOQ solutions: avoiding stockouts or carrying excess inventories. We find that iterations of the task quickly diminish the probability of making such choices and, surprisingly, imposing high cognitive loads doesn't affect these probabilities.

Our study is one of the first to experimentally examine a stationary limited horizon EOQ model. We found two previous studies that examine infinite horizon EOQ environments. The EOQ is one of the three environments Stangl and Thonemann (2017) consider in their behavioural study of inventory decision-making under two common alternative frames of performance measurement: inventory turnover and the number of days of inventory held. The former leads managers to over-value inventory reductions relative to the latter. K.-Y. Chen and Wu (2017) examine learning in an infinite EOQ environment in which there is varying inventory ordering and holding costs. The experiment consists of fifty rounds of such inventory decisions. For the first fifteen rounds operational costs were constant, and they varied during the last thirty-five rounds. Their result shows that learning occurs over rounds, and participants learn much faster about the optimal choice under stable environment than under changing environment. Suboptimal decisions tend not to be repeated with deterministic feedbacks. It is important to note that their participants' choice sets are even more restricted than those of our Zero Only treatment. Participants are required to choose from an EOQ restricted choice set whose elements are the number of weeks, their periodicity of demand, of inventory ordered each time inventory is depleted. Thus, their policy choice set consists only of EOQ policies with fixed EOQ cycle lengths. The feedback K.-Y. Chen and Wu (2017) provide participants is the average operational costs generated per week by their EOQ cycle length choice, and participants' reward metrics are the sum of their average weekly performances. While we provide a monthly reported feedback on each decision made, participants experience and collect rewards on a month-to-month basis, which will vary from months when inventory is ordered to those when it is not.

3.2 Experiment

3.2.1 Inventory decision task

As discussed in section 2.2, the core decision-making part of our experiment is an inventory management task, where the participant has to manage 'S-store' that sales coffee makers. Participants were asked to make monthly ordering decisions for a consecutive of six years.

Table 2 compares the efficiency of those EOQ constant policies with no stockouts incurred in any month, or any positive closing inventories in storage by the end of month 12. The optimal inventory policy has 100% efficiency in terms of its annual profits generated. Notice that EOQ constant 40 and EOQ constant 20 can also generate over 93% of the optimal annual profits. Given the minimal loss incurred by adopting these policies, we define EOQ constant 40 and EOQ constant 20 as "near optimal" performance.

Table 2: Alternative EOQ constant policy performance efficiency table

EOQ constant	Orders per year	Annual profit	Efficiency
120	1	75	15.63%
60	2	390	81.25%
40	3	465	96.88%
30	4	480	100.00%
20	6	450	93.75%
10	12	240	50.00%

3.2.2 Experimental design

We implement a 2 x 2 experimental design with two treatment variables that exogenously impose cognitive stress. We adopted a between-subject design where a participant randomly takes part in one of the four treatments. The first treatment variable is the feasible set of inventory policies a participant can follow. The first category is called "Unrestricted", where a participant can choose any quantity they wish each month as long as the quantity does not exceed 120. The second category is called "Zero Only", where participants are restricted to ordering only once the inventory level is zero. We expect that the larger set of alternatives in the unrestricted category will present participants with a more difficult learning task.

The second treatment variable is the level of exogenously imposed cognitive load. We introduce a concurrent task that competes for subjects' cognitive recourse. In the "Low" cognitive load category subjects complete the inventory management task without distractions. In the "High" cognitive load category we introduce an incentivised PIN task that requires subjects to memorize in their short-term memory while completing the inventory management task. At the beginning of each year, subjects will be given 15 seconds to memorize a random 6-digit PIN code. The PIN code is case sensitive, consisting of numbers, lower and upper case letters.[1] At the end of the year, subjects will be asked to recall the PIN code. Subjects only have one attempt to enter the PIN code; successful recall will unload an extra reward of P300. We expect actively take part in the memorization task reduces subjects' short-term working memory to complete

[1]The PIN is the same for all participants across each year to ensure control.

the inventory management task, consequently, diminish the ability of learning.

At the end of the experiment we collected information in a short questionnaire about the participants' age, gender, level of education, major and math skill level. Table 3 summarizes our experimental design and provides summary statistics on the demographics of the participants. We designate treatment cells by the word pairs x-y, where x is feasible set of policies category and y is category of the cognitive load.

Table 3: Summary of the demographic information of participants for each treatment

Treatment cell	Participants	Average age	Male	Postgrad	STEM subjects[1]	Average math level[2]
Unrestricted-Low	41	25	34%	49%	37%	3.68
Unrestricted-High	41	25	37%	44%	56%	3.20
Zero Only-Low	39	25	23%	47%	34%	3.26
Zero Only-High	36	28	50%	56%	28%	3.53

[1] STEM subjects include Engineering & Technology, Life Sciences & Medicine and Natural Sciences. Non-STEM subjects include Arts & Humanities and Social Sciences & Management.
[2] Math Level was self-assessed, and was categorised into 6 levels. 1 = "Below GCSE", 2 = "GCSE", 3 = "A Level", 4 = "Undergraduate", 5 = "Postgraduate", 6 = "Above Postgraduate". Note that GCSE (General Certificate of Secondary Education) is an academic qualification in a specific subject typically taken by school students aged 14-16 of the UK (except Scotland), at a level below A level.

3.2.3 Experimental procedures

We conducted seven experimental sessions at Newcastle University Business School experimental economics laboratory during May and July 2017. The participants were recruited from the subject pool of the Behavioural Economics Northeast Cluster via invitation emails. In total, 162 subjects participated in the experiment[2]. All participants were students from Newcastle University except for 3 who were from Northumbria University.

Each session lasted no more than 60 minutes, with strict procedures to limit the access to any aides that would provide assistance in calculations or remembering PIN codes. Participants were signed in individually and instructed to leave their personal belongings, including any writing instruments, in the reception area before shown to a computer desk placed in a privacy carrel. Each participant was then provided with a pen and two copies of an informed consent document. They were asked to read and signed on the documents if they wished to continue their participation. An experimenter then collected the pen and signed forms. Participants were informed thereafter that no electronic devices - such as mobile phones, calculator, smart watches, etc. - could be used until their session was completed. The information displayed on their computer monitor is private and the decisions are made individually. They were further instructed that the experimental tasks were fully computerized and they would complete the

[2] We excluded 5 participants from our data analysis and the participant counts are given in Table 3. One participant, in the Zero Only-Low treatment, always submitted the random slider starting position when inventory reached zero. Two other participants, in the Zero Only-High treatment, grossly took advantage of the limited liability rule. The final two excluded participants attended the last session and demonstrated behaviour that they had been briefed about the content of the experiment; they clicked through the instructions without reading them and subsequently provided the solution $Q^* = 30$ for all years - even though this was not optimal for the practice year.

rest of the experiment only using their mouse. Prior to participants entering the laboratory, all computer keyboards were concealed under a thick opaque cover. This was to done to eliminate any chance to write down the PIN codes. These measures were taken in all sessions to provide control between High and Low cognitive load treatments. The experiment was programmed and conducted using a self-contained program developed in oTree (D. L. Chen, Schonger, & Wickens, 2016). Access to other programmes on the computer was restricted to eliminate many of the tools participants commonly used to perform mathematical calculations. Participants can ask questions to an experimenter at any point during the experiment session.

Once instructed to start by the experimenter, participants started to read through the instructions[3] at their own pace. After reading the instructions, participants were asked to complete 7 multiple choice questions designed to ensure that they understand the calculation of costs and profits. Participants who missed more than 2 correct answers had to review the questions with one of the experimenters before proceeding to the inventory management tasks.

Participants then took part in the six-year decision task sequence, followed by a short post-experiment survey which collected demographic information. Year 0 was a practice round which used an alternative set of cost parameters[4] from those of Years 1 through 5, and the performance in this task did not affect a participant's total earnings. The purpose of the practice year is to give the participants an opportunity to get familiar with the software and decision tasks. Participants used a moving slider range from 0 to 120 to select the ordering quantity. The initial point on the slider starts from a random point each month. In the EOQ treatment, the starting point on the slider is greyed out when the opening inventory level is positive. To help participants with calculations of the task, we provide basic formulas, key information regarding the current month, and a monthly record table that calculates the sales revenue, costs, and profits on the decision screen, as well as an annual profit table that displays the profits made in each year from Years 1 through 5.[5] For participants who experienced the High cognitive load treatment, we provided an opportunity to practice the PIN task in the practice year.

Participants then completed the Years 1 through 5 decision tasks. At the end of the experiment they were paid £5 show-up fee and their accumulated earnings from the decision tasks, converted to pounds at the conversion rate of ₽300 = £1. There was limited liabilities; to ensure that no one leaves the experiment with a payment less than £5 or affected by a large negative earnings made in a particular year, any negative profits made in a year will be treated as 0 earnings.[6] The average earning in the experiment is £13.37 per participant, including the show-up fee.[7] The payments were handed out to participants in sealed envelopes, so that the amounts received in the experiment were not revealed to other participants.

One last important aspect of the experiment is that we require the participants to spend a fixed amount of time in completing the inventory management task for a year. We required that

[3]In the supplement appendix, we provide a complete set of instructions.

[4]In the practice year the order costs were ₽45 and the holding costs were ₽0.5.

[5]We provide screen captures of these interfaces in the supplement appendix.

[6]This limited liability only affected the earnings of 5 participants in 5 different years.

[7]The average earnings of Low treatment (without PIN task) was £11.58 per participant, while the average earnings of High treatment (with PIN task) was £15.23 per participant.

a participant spend exactly 4 minutes completing each task in Years 1 through 5. This was designed to prevent participants from racing through the monthly decisions in order to reduce the cognitive cost of remembering their PIN. If a participant completed their twelve monthly decisions early they could not advance to the next period (or enter the PIN) until the four minutes expired. The programme would jump to a wait page and review the results of the current year. If the participant failed to complete the 12 monthly decision tasks before the time expired, the computer program will automatically execute the remaining month sales with the existing inventory in stock.

3.3 Hypotheses

Our motivation of the treatment variables leads us to the following hypotheses. Naturally, better performance leads to greater average annual earnings and is indicated by greater percentage of participants adopting optimal (near-optimal) inventories. An additional concurrent PIN task that competes for subjects' cognitive resources reduces the capacity of short-term working memory. Therefore, subjects' annual profits and the ability to perform the EOQ optimal policy would be negatively affected.

Hypothesis 1. *Participants perform better in the Zero Only-Low treatment than the Zero Only-High treatment. This will be reflected in two measures: average annual earnings and the percentage of participants who adopt optimal (near-optimal) inventories.*

Hypothesis 2. *Participants perform better in the Unrestricted-Low treatment than the Unrestricted-High treatment. This will be reflected in two measures: average annual earnings and the percentage of participants who adopt optimal (near-optimal) inventories.*

The set of inventory policies in the Unrestricted is much larger than and only adds suboptimal alternatives to the Zero Only restricted set of policy choices. The reduced focalness of EOQ strategies and greatly complicated participants' choice sets in the Unrestricted treatments.

Hypothesis 3. *Participants perform better in the Zero Only-Low treatment than the Unrestricted-Low treatment. This will be reflected in two measures: average annual earnings and the percentage of participants who adopt optimal (near-optimal) inventories.*

Hypothesis 4. *Participants perform better in the Zero Only-High treatment than the Unrestricted-High treatment.This will be reflected in two measures: average annual earnings and the percentage of participants who adopt optimal (near-optimal) inventories.*

3.4 Results

We evaluate the treatment effects of restricted inventory policy choice sets and increased cognitive load by considering their impacts upon participant's earnings in the inventory management tasks, the propensity to choose optimal inventory policies, and then the efficacy of the PIN task and whether performance in that task is correlated with inventory performance.

3.4.1 Annual inventory profits

We test the differences in average annual profit for different treatment groups using two-sided t-tests and non-parametric Wilcoxon rank-sum tests. As participants have the best performance with the intervention and without the cognitive load, we use Zero Only-Low treatment as a reference point to demonstrate profit loss resulting from the presence of more complicated choice sets and a shock to their cognitive load.

We report the average annual profits made by subjects by treatment groups in Table 4 Panel A. The results indicate that both the absence of the intervention and shocking their cognitive load negatively impact average annual profits both statistically and economically. More complicated policy choices cause more profit loss than High cognitive load. The test results of the differences in average annual profit for different treatment groups are reported in Table 4 Panel B. We find that exogenously imposed cognitive load significantly reduced the average annual profits in the Zero Only treatment; while the reduction is not statistically significant in the Unrestricted treatment when the performance are already sub-optimal in general. Thus, the results support Hypothesis 1. Further, we find stronger evidence in support of Hypothesis 3 and 4. Limiting participant's choices to EOQ restricted policies leads to statistically greater average earnings in both Low and High cognitive load settings.

Table 4: Average annual profits by treatment and hypotheses tests for differences in average annual earnings

Panel A: Annual profits by treatment

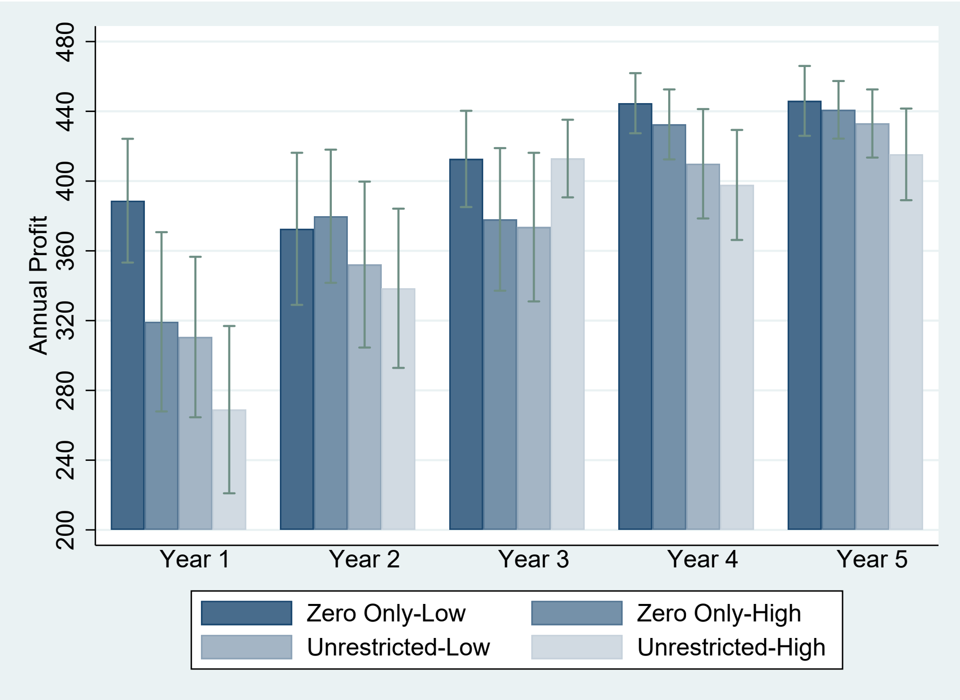
	Unrestricted-Low	Unrestricted-High	Zero Only-Low	Zero Only-High
Average	375.85	366.70	412.94	390.10
Stand. Dev.	129.38	126.87	97.35	113.78

Panel B: Hypotheses tests for differences in average annual profits (p-values reported)

Treatment Comparison	Difference	Profit loss (%)	Two-sided t-tests	Wilcoxon rank-sum
Zero Only vs Unrestricted	30.71	7.64%	0.000	0.001
Low vs High	16.29	4.14%	0.055	0.003
Zero Only-Low vs Zero Only-High	22.84	5.53%	0.038	0.012
Unrestricted-Low vs Unrestricted-High	9.15	2.43%	0.470	0.124
Zero Only-Low vs Unrestricted-Low	37.09	8.98%	0.001	0.012
Zero Only-High vs Unrestricted-High	23.40	6.00%	0.059	0.052

Figure 4 presents the average profits made in Year 1 through Year 5 by treatment groups. It provides a disaggregated view of the average annual profits permits insight into learning over time and how our treatments impact it. This figure has several prominent features that give us insight into the hypotheses results on average profits levels. First, subjects in the Zero Only-Low treatment achieved high initial annual profits, around 80% of the optimal profit in Year 1; High cognitive load and Unrestricted policy choice sets both cause the greatest negative performance impact in Year 1. Second, performance gains are mostly achieved in Years 1 through 3. Third, average annual profits are around 90% of the possible earnings in the last two years; except for the Unrestricted-High treatment which is around 5-10% lower.

Figure 4: Annual Profits over individual Years and by treatment: Averages and 95% confidence intervals



We quantify our results by conducting a series of dummy variable linear regressions using random effects estimators and cluster standard errors at the level of the individuals. We report these results in Table 5. In model (1) we regress annual profits on dummy variables from Year 1 through 4. The average annual profits made in Year 5 serves as the base category. In model (2) we introduce dummy variables for the Unrestricted and High treatment categories. In this case the constant reflects the average profit level for Year 5 in the Zero Only-Low treatment; and the Year 1 through 4 dummy variable coefficients reflect the average annual profits across participants in the Zero Only-Low treatment. In model (3), we add interaction dummy variables for the Unrestricted and High treatment categories for the joint effect of imposing both treatments on annual profits. In model (4), we add individual characteristic dummy variables to examine individual differences.

From the regression results we can conclude that the learning in terms of average annual profits earned are significant in Year 1 through Year 3. Our treatment effect of Unrestricted and High are mostly generated because of the impact on subjects' performance in Year 1, when the participants face the inventory decision problem for the first time, as suggested by their individually significant coefficients in models (2) and (3). Further, conduct a Chow test, for which the null is model (1) versus the alternative of model (2), i.e. the joint differences of the two treatments are significant. The result supports the successful implement of the two treatments (χ^2 (10, N=785) = 27.32, p = 0.002). We then conduct a second Chow test between model (2) and (3) for the joint significance of the two treatments. The result suggests there is no effect of simultaneously imposing both the Unrestricted and High treatments (χ^2 (5, N=785) = 8.73, p = 0.120). Our results in model (4) show that undergraduate students earn slightly more in terms of profits, while male participants perform better than female.

Our analyses of annual profits leads us to our first set of results.

Result 1. *Reducing the participants' policy choice sets to EOQ restricted ones leads to higher profits. However, these gains predominantly occur in Year 1 - when the participants face the inventory decision problem for the first time.*

Result 2. *Exogenously increasing participants' cognitive load leads to lower profits. However, these losses predominantly occur in Year 1 - when the participants face the inventory decision problem for the first time.*

Result 3. *There is no super- or sub-additive effect of simultaneously exposing participants to the Unrestricted and High treatment categories.*

Table 5: Dummy variable regressions for annual profit. (n=785)

	(1) Annual Profit	(2) Annual Profit	(3) Annual Profit	(4) Annual Profit
Year 1	-112.27***	-67.35***	-57.22***	-58.53***
	(11.24)	(16.28)	(16.88)	(17.32)
Unrestricted*Year 1		-45.45**	-65.22**	-63.91**
		(21.97)	(26.77)	(27.09)
High*Year 1		-43.20*	-64.31**	-63.00**
		(22.17)	(30.23)	(30.54)
Unrestricted*High*Year 1			40.40	39.09
			(44.19)	(44.46)
Year 2	-73.39***	-71.29***	-73.36***	-75.82***
	(10.02)	(16.75)	(19.05)	(19.44)
Unrestricted*Year 2		-11.56	-7.53	-5.07
		(20.05)	(29.81)	(30.10)
High*Year 2		8.04	12.35	14.80
		(20.13)	(26.73)	(27.04)
Unrestricted*High*Year 2			-8.24	-10.69
			(40.13)	(40.40)
Year 3	-38.81***	-54.99***	-33.27**	-31.07**
	(7.78)	(12.53)	(13.37)	(13.57)
Unrestricted*Year 3		16.24	-26.15	-28.35
		(15.78)	(22.69)	(22.85)
High*Year 3		15.70	-29.56	-31.77
		(15.65)	(23.53)	(23.69)
Unrestricted*High*Year 3			86.59***	88.79***
			(30.82)	(30.97)
Year 4	-12.85***	-4.49	-1.33	-1.18
	(4.61)	(5.26)	(3.74)	(3.85)
Unrestricted*Year 4		-15.59*	-21.75*	-21.90*
		(9.03)	(12.16)	(12.23)
High*Year 4		-0.45	-7.03	-7.18
		(9.25)	(9.26)	(9.32)
Unrestricted*High*Year 4			12.59	12.74
			(18.12)	(18.19)
Unrestricted		-19.12*	-12.96	-14.30
		(10.31)	(13.87)	(15.17)
High		-11.70	-5.13	-7.18
		(10.45)	(12.83)	(14.83)
Unrestricted*High			-12.58	-15.31
			(20.71)	(22.68)
Male				15.04
				(11.85)
Postgrad				-14.88
				(12.03)
STEM				12.18
				(11.47)
Math Level				-2.69
				(5.13)
Constant	433.40***	449.13***	445.97***	454.89***
	(5.28)	(8.70)	(9.92)	(16.35)
chi2	183.28	210.84	228.77	239.23
p	0.00	0.00	0.00	0.00

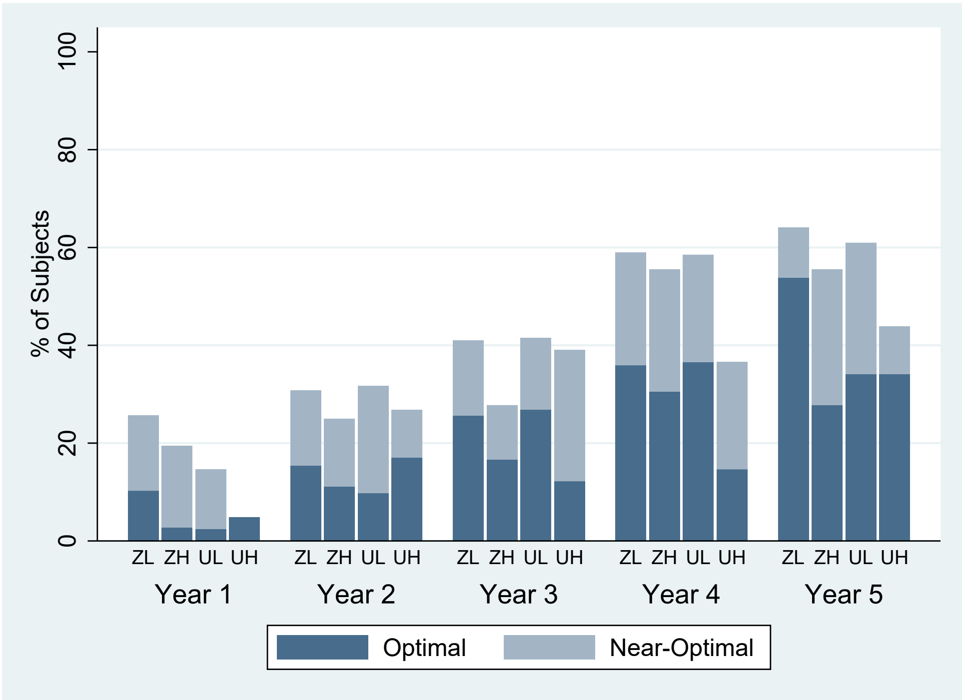
Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.4.2 Inventory management policy choices

We turn our analysis towards the inventory policy choices of participants. We analyse each participant's inventory policy choices on an annual basis. For whether $Q_{i,a}$ is optimal, if $Q^* = 30$;

or $Q_{i,a}$ is near optimal, if the annual order decisions are any combinations of EOQ policies of 20, 30 and 40. Figure 5 shows the percentage of participants following optimal and near optimal strategies in each treatments. In general, the percentage of subjects adopting optimal or near optimal policy is higher in the Low treatments than in the High treatments for both Zero Only and Unrestricted in all five years. In particular, subjects in the Zero Only - Low treatment perform the best in finding the optimal or near optimal policy. We can see a general trend in the increase of adopting optimal and near optimal policy from Year 1 through Year 4 in all treatments.

Figure 5: Stacked graph of the percentage of participants following optimal and near-optimal EOQ constant strategies: by Year and treatment



Next, we investigate whether the implementations of Zero Only and High load treatments are influencing the probability of subjects to adopt EOQ optimal or near optimal policy. The Logit regression results are presented in Table 6.

First, as indicated in the last row of the table, the average levels of the estimated probability of choosing an optimal policy are lower than the probability of choosing a near optimal policy. Second, the coefficients on Years are positive and significant indicating there is a significant trend in learning to make EOQ decisions from Years 1 through 5. Third, the coefficients on High are negative and significant, which is indicating that the probabilities of choosing optimal or near optimal policies are reduced by the implementation of exogenous cognitive load. Whereas the support from the Zero Only treatment did not improve the probabilities of choosing optimal or near optimal policies, as indicating by the insignificant coefficients on Unrestricted.

Result 4. *There is a trend in all treatments for increasing use of optimal and near-optimal policies across Years.*

Result 5. *High cognitive load leads to lower percentage use of these policies for both Zero Only*

Table 6: Logit regression on the probability of choosing optimal or near optimal policy. (n=785)

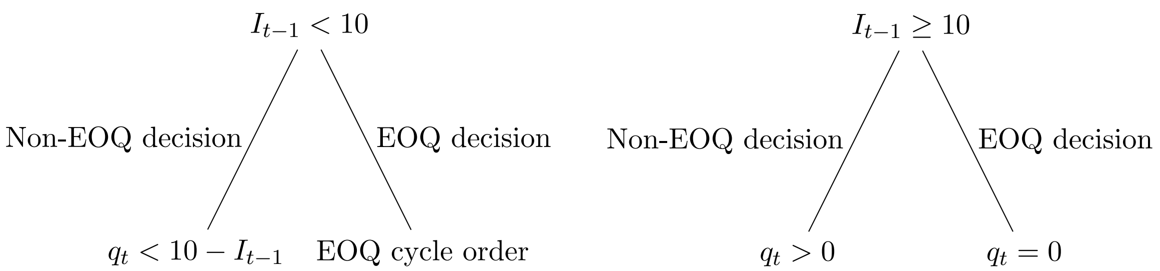
	(1) Optimal	(2) Optimal	(3) Near-optimal	(4) Near-optimal
Year	0.529***	0.536***	0.468***	0.474***
	(0.055)	(0.056)	(0.048)	(0.049)
High		-0.516*		-0.443*
		(0.269)		(0.236)
Unrestricted		-0.249		-0.212
		(0.262)		(0.234)
Constant	-3.059***	-2.721***	-1.940***	-1.639***
	(0.236)	(0.275)	(0.188)	(0.246)
χ^2	93.58***	90.46***	94.72***	96.38***
Pr	0.211	0.211	0.381	0.381

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

and Unrestricted in all five Years.

Finally, as in our experiment design, most part of the Non-EOQ decision was eliminated in the EOQ treatments; then the key interest is whether imposing High cognitive load leads to higher probabilities of Non-EOQ decisions. There are two key logical motivations for making decisions under the EOQ policy, avoiding stock out[8] - thus do not loss potential profits; and only place an order when current inventory depleted - thus avoiding excess holding costs. We present the decision process for monthly choices as a decision tree in Figure 6, where the decision is made upon whether the closing inventory of previous month is greater or less than 10.

Figure 6: The branching decision process.



(a) Decision process when the closing inventory of previous period is strictly less than ten.

(b) Decision process when the closing inventory of previous period is greater than ten.

We use a simple Logit regression to quantify the probability of making Non-EOQ decisions as a function of time, accumulated Non-EOQ ordering habit, and whether it is a High cognitive

[8] We recognise that with our setting, especially in later months, it may be more profitable to suffer a stockout when the open inventory is not too far short from the demand. For instance, in month 9 the optimal dynamic solution is to order 40. However, if open inventory is above 6 units in month 9, it would be more profitable to suffer a stockout and wait until period 10 to order 30. This may lead to a situation in Unrestricted treatment, in which participant deliberately wait out a stockout. However, out of 4920 observations from Unrestricted treatment, such situation never occurred.

load treatment. In any one of the decision rounds, we define

$$NonEOQ_{i,r} = \begin{cases} 1 & \text{if } q_{i,r} \text{ is not an EOQ decision in decision round } r\text{, and} \\ 0 & \text{otherwise.} \end{cases}$$

where $r \in \{1, 2, \ldots, 60\}$.

We estimate a set of Logit regressions on the probability of a subject makes Non-EOQ decision for two conditions; one when the closing inventory of the previous month is strictly less than 10 and one when it is greater or equal to 10. Under both conditions, our estimation is based on the following specification

$$Pr(NonEOQ_{i,r} = 1) = F(\beta_0 + \beta_1 Year_r + \beta_2 Month_r + \beta_3 High + \beta_4 NonEOQACC_{i,r-1}).$$

where F is the cumulative logistic distribution function. We introduce variables to capture the habit of making Non-EOQ decisions in model (3) and (6). $NonEOQACC_{i,r-1}$ is the running count of the total number of rounds the subject i has deviated from the EOQ decision up to the round $r-1$. The Logit regression results are presented in Table 7: Panel A for the case $I_{t-1} < 10$ and Panel B for the case $I_{t-1} \geq 10$. Under the former condition, when the closing inventory of the previous month is less then 10, Non-EOQ decisions generate stockouts and miss opportunity to achieve potential sells. In our experiment, subjects in the Unrestricted treatment have the opportunity to make such decisions. While in the Zero Only treatment, subjects may have positive closing inventory of the previous month that is less than 10 but are not allowed to place an order. We exclude these observations (138 out of 4500 observations), as they were not made by active choice. There is another possibility in the Zero Only treatment when subjects have none closing inventory in the previous month but choose to order less than 10. We include these observations (40 out of 4500 observations) in our analysis. Under the later condition, when the closing inventory of the previous month is greater or equal to 10, Non-EOQ decisions generate excess holding costs to keep orders in stock. In our experiment, only subjects in the Unrestricted treatment are allowed to order a positive amount under such conditions. Therefore, observations from the Zero Only treatment are excluded from the analysis.

First, from the regression results we have large negative values of the estimated constant, thus, the average levels of the estimated probability of making a Non-EOQ decision are very small as indicated in the last row of the table. Second, the coefficients on Years are negative and significant indicating there is a significant trend in learning to make EOQ decisions from Year 1 through Year 5. The estimated coefficients on Months are significant with smaller magnitudes, and have opposite signs under the two conditions. This suggests that stockouts are more likely to happen later in a year while having excess inventory in stock is less likely to happen later in a year. Third, the estimated coefficients on the accumulated Non-EOQ decisions are positive and significant captures the individual differences in adopting the EOQ ordering logic. Lastly, the effect of imposing high cognitive load on deviating from the EOQ decisions is not statistically significant. Therefore, the performance differences in adopting EOQ policies must come from the profitable level of EOQ policies taken in the High cognitive load treatments. Overall we interpret this evidence that providing the more complicated choice set does lead to some Non-EOQ decisions, but these choices diminish with experience.

Table 7: Logit regression on the probability of deviating from an EOQ decision

	Panel A: $I_{t-1} < 10$			Panel B: $I_{t-1} \geq 10$		
$NonEOQ_{i,r}$	(1)	(2)	(3)	(4)	(5)	(6)
$Year_r$	-0.498***	-0.500***	-0.690***	-0.308***	-0.309***	-0.608***
	(0.126)	(0.128)	(0.160)	(0.096)	(0.097)	(0.134)
$Month_r$	0.194***	0.194***	0.153***	-0.115***	-0.114***	-0.141***
	(0.047)	(0.047)	(0.054)	(0.027)	(0.027)	(0.029)
High		0.255	0.216		-0.404	-0.193
		(0.425)	(0.311)		(0.376)	(0.296)
$NonEOQACC_{i,r-1}$			0.288***			0.307***
			(0.039)			(0.040)
Constant	-3.379***	-3.507***	-3.175***	-1.754***	-1.573***	-1.311***
	(0.583)	(0.630)	(0.690)	(0.338)	(0.397)	(0.430)
N	3032	3032	2875	3286	3286	3286
χ^2	34.10***	36.06***	97.89***	22.66***	22.90***	75.76***
$Pr(NonEOQ_{i,r}) = 1$	0.034	0.034	0.037	0.035	0.035	0.030

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

3.4.3 Efficacy of the PIN reward procedure

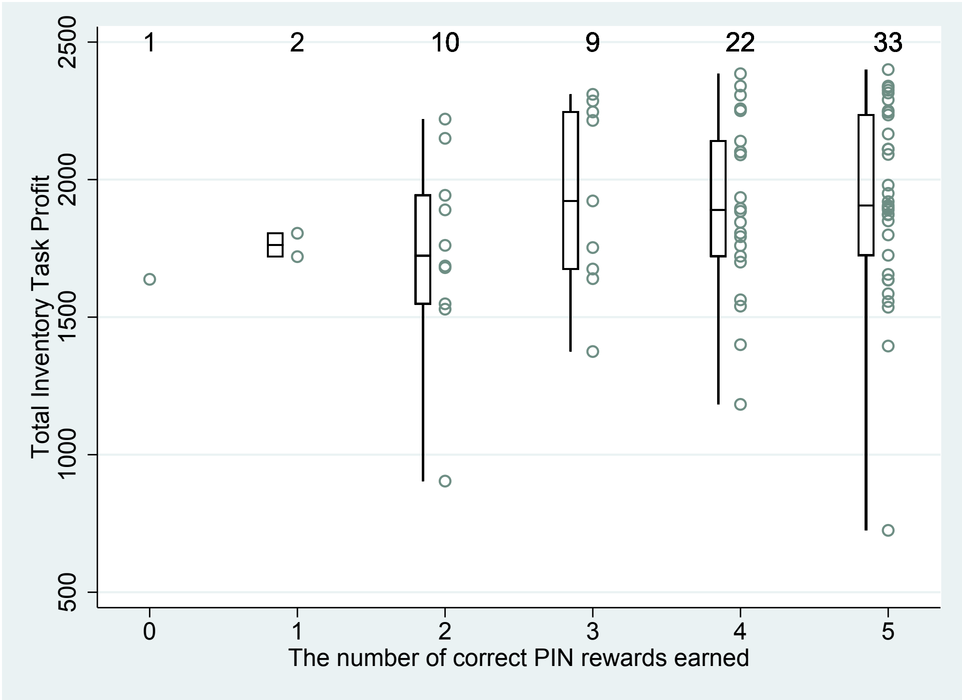
Next, we evaluate the efficacy of the PIN reward task that exogenously imposed cognitive load on the High cognitive load treatments. In our pilot experiment sessions, we tested the treatment effect with both 6-digit PIN task and 7-digit PIN task given 20 seconds for subjects to memorize. When the PIN reward task is too simple, participants always collect the reward without taking up a sufficient part of the available cognitive resource that actually affects the performance on the inventory management task. When the PIN reward task is too difficult, participants could either decide to forgo the mental costs of trying to commit the PIN to short term memory or forgo effort in the inventory management tasks. A second concern is that raw intelligence is an omitted variable in our analysis which would manifest itself in a strong positive correlation between a participant's performances in the PIN reward and the inventory management task.

Therefore, we adjusted our design to balance the effort spent on the two concurrent tasks. In the High cognitive load treatments, participants were given 15 seconds to memorize a 6-digit PIN code. The success rates of correctly recall the PIN is at 49% for the Unrestricted treatment and 72% for the Zero Only treatment; the difference between the two treatments is significant under both t-test ($p = 0.037$) and non-parametric Wilcoxon rank-sum test ($p = 0.038$). Further, participants indicate in the post-experiment survey that they cared equally about the PIN task and the inventory management task (mean equal to 2.87 on a 1-5 scale) with no statistical differences between the two treatments (t-test, $p = 0.355$; Wilcoxon rank-sum test, $p = 0.511$). To clarify, our result is not driven by subjects' rational allocation of attention. Firstly, as discussed in the introduction section of the current chapter, previous studies proposed that the PIN task as the instrument to impose cognitive stress in laboratory experiments. We have also discussed about our design of avoiding the PIN code being too simple or too difficult. Further, we will provide statistical analysis around the independence of PIN and inventory management task performance in the following paragraphs.

We present subjects' total profits made in the inventory management task conditional on the number of PIN rewards earned in Figure 7. The numbers on the top are the number of subjects who earned the corresponding to number of PIN rewards. First, we can observe that only 3 out of 77 subjects earned less than two PIN rewards while 33 out of 77 subjects collected all five PIN rewards. Second, there is no clustering of poor inventory management performance on either high or low numbers of PIN rewards earned. This is indicating that the raw intelligence of participants is irrelevant to the performance on the PIN task and subjects are balancing the effort spent between the two concurrent tasks. Third, there is little evident difference in the conditional means of total profits; suggesting the PIN and inventory management tasks performance are independent.

We quantify the evidence of the independence of PIN and inventory management task performance by statistically measuring their correlation and testing its statistical significance. Table 8 reports the test results of the independence of the performance on the PIN reward task and the inventory management task where the null hypotheses assume that the correlation is zero. The left portion of the table presents the correlation between the success of a PIN reward task in the year and the corresponding annual inventory profit. We find mixed evidence that

Figure 7: Participants' total inventory management task profits conditional on the number of PIN rewards earned and the corresponding whisker plots for the 50, 75, and 95% quantiles.



the correlation is not significantly from zero in all five years separately, but highly significant positive correlation when we pool all five years together. However, this part of the analysis does not allow for differences in terms of subjects' performances on the PIN task. To address the problem, we test the correlations between the total numbers of PIN reward earned by a subject and the subject's annual profits as well as the subject's total profits in the inventory management task. The results are reported in the right portion of the table. From the tests results we can conclude that there is no significant correlation between the two tasks.

Table 8: Spearman correlations between PIN reward earned in a Year and the corresponding Inventory task profit; Spearman and Pearson Rank correlations between a participant's total number of earned PIN rewards and their Inventory task profits

		PIN reward eared in Year a	Number of PIN reward earned	
		Spearman Rank Corr.	Pearson Corr.	Spearman Rank Corr.
Annual Profit	Year 1	0.08 (0.512)	0.13 (0.248)	0.11 (0.324)
	Year 2	0.12 (0.315)	0.08 (0.507)	0.11 (0.338)
	Year 3	0.21 (0.072)	0.10 (0.391)	0.04 (0.750)
	Year 4	0.09 (0.459)	0.15 (0.182)	0.04 (0.725)
	Year 5	0.14 (0.226)	0.10 (0.379)	0.14 (0.218)
	All Years	0.18 (0.001)	N/A N/A	N/A N/A
Total Profit		N/A N/A	0.179 (0.120)	0.182 (0.114)

1. The p-values of the respective tests are reported in the parenthesis.
2. We don't report the correlations for Total Profit in column three because the calculation will include multiple repetitions of a participant's total inventory profit.
3. We don't report the correlations for all Years in columns for and five because the calculation will include multiple repetitions of a participant's total number of PIN rewards.

3.4.4 Gender-multitasking

Buser and Peter (2012) define multitasking as a 'task-switching' between multiple ongoing tasks from time to time, which can be applied in our treatments with the PIN task. In our High load treatments that involve the PIN task, participants need to memorize the PIN as well as making monthly ordering decisions. Popular best-selling books advertise the stereotype that women are better at multitasking (Pease & Pease, 2001, 2003). However, experimental results from Buser and Peter (2012) do not support such stereotype. In our experiment, actively maintain the PIN to be recalled later requires retrievals to refresh the memory. Participants switch the attention from completing the inventory task to refresh the PIN code for several number of times during the time constrain. Our experiment is an opportunity to evaluate whether women are better at multitasking conjecture in an inventory management context.

We are interested in the differences in average annual profit between male and female participants in the treatments involving the PIN tasks. Among the participants who needed to memorize the PINs as well as making monthly ordering decisions, the average annual profits of female and male are 362.95 and 397.22 respectively. Surprisingly, women are not better at multitasking as stereotype suggests. This difference of 34.26 is statistically significant according to both a t-test ($p = 0.006$) and a Wilcoxon rank-sum test ($p = 0.000$). Therefore, we have no evidence to support that women are better at multitasking than men when there is an concurrent task in conjunction with inventory management task.

3.5 Managerial Application

Our research provides managerial insights for a set of practical problems. The actual practice of EOQ inventory management model is prevalent in wholesale distributions for managing durable goods. EOQ is applied to an extensive type of products including automobiles, furniture, home appliances, and sports equipment. It is particularly used when managers are responsible for managing multiple product lines. For example, we have conducted interviews with inventory managers from four national and international firms based in Guangdong and Jiangsu, China. The firms produce durable goods such as furniture, clothing and household hardware. In their daily operations, each of the inventory managers is responsible for managing multiple product lines. Also, products are often categorized into different groups based on their condition (e.g. finished or semi-finished, fully paid or paid in deposit). It is cognitively demanding for managers to treat multiple product lines differently while coordinating with other departments. They have reported that a major source of performance pressure comes from the common assignment of new product lines. We speculate that these pressures consume their cognitive resources as they need to process rates of demand, ordering time frames, holding costs et cetera.

During our interviews, the managers have testified that the ERP system is the predominant tool they use to support managing multiple product lines, where EOQ model is the built-in calculation for ERP software packages (Oracle, 2018). These modules automate the tracking of inventories and sales, as well as recommend order timing and size based upon EOQ solutions. Managers rely on the system's records to keep track of the inventories in stock, sales and ordering activities. Systems are good at providing real time solutions and yielding impressive profits improvement for companies when dealing with the same environment repeatedly. However, systems cannot act as a fully replacement for human oversight. Managers are responsible for evaluating and making actual ordering decisions to overwrite the deterministic solution from the system. Human talent is critical to supply chain success especially in more complex situations, they also noted that they retain, and often exercise, the ability to deviate from the system's recommendations. In the case where firms launch new product lines, inventory managers are also responsible for updating the inventory account inputs in the system every month accordingly. Thus, human biases and judgement errors also create potential inefficiencies.

3.6 Conclusion

In Chapter 3, we conducted an experimental research to assess the effect of cognitive stress on inventory management decisions using the EOQ inventory management model. We manipulated the level of cognitive stress in both directions in our experimental design. The first treatment exogenously imposed cognitive stress from a PIN task that competes for the participants' short-term memory resources, and the second treatment reduces cognitive stress by introducing an intervention that limits the complexity of the inventory policy choice set.

Our results show that increases in cognitive load negatively impact participants' performance. Participants earn less when there is a competing task or when the policy choice set is not restricted; restricting the potential policy choices reduces the impact of High cognitive load. However, these negative impacts occur predominantly when participants first face the inventory decision problem. We observe there is a trend that participants learned to adopt near optimal

EOQ policies in general. We note that only in the Zero only-Low treatment cell do we observe the majority of participants eventually learn to use the optimal EOQ policy. The restriction of managers to only place orders when inventories are exhausted and the alleviation of the competing task improved the chance for inexperienced decision makers to reach the optimal inventory policy. It should be noted that these performance differences and suboptimal choices largely occur in the first three iterations of our environment. We find that iterations of the task quickly diminish the probability of making suboptimal choices that lead to stockouts and other choices that lead to carrying excess inventories in both High and Low cognitive load treatments. The results also imply cognitive stress treatment only affects participants' performance under Zero Only condition, but there's no evidence to support such difference under the Unrestricted condition where the performance is already impaired by not getting the support from the EOQ system. The real word implication is to show the importance of alleviating inventory managers' cognitive stress from managing multi-product lines and allocation new products even though they are supported by the EOQ system.

The present research aims to assess the impact of cognitive stress on inventory management performance in a static experimental setting, and has developed a positive intervention to minimize the negative impact. Higher exogenous cognitive load has more severe impact on participants' performance when policy choice sets are not restricted to EOQ type of behaviours. However, we have observed some individual differences in terms of gender and level of education, which points to an interesting direction for our next step of research. We believe our current experimental design provides a good foundation to explore further topics in inventory management for durable goods.

In the next chapter we are going to study whether individuals' level of cognitive ability would be a good indicator their performance in supply chain management.

4 Cognitive Ability

4.1 Introduction

Higher cognitive ability is explicitly screening by employees in the job recruiting process across a wide range of industries, particularly in quantitatively oriented positions. According to Harper (2008), 85% of the FTSE 100 companies are using psychometric test for recruitment in which a majority part consists of cognitive ability tests. Level of cognitive abilities has been proven to be a good predictor for workplace performance of new hire (Newman & Lyon, 2009; Schmidt & Hunter, 1998). Job candidates with higher cognitive abilities achieve higher workplace performance. However, recruiting on cognitive ability may also lead to a less diverse workforce, especially gender difference in cognitive tests performance has also been observed in several independent investigations (Campitelli & Gerrans, 2014; Pennycook et al., 2016; Primi et al., 2016).

In behavioural operations management, existing research studies have identified that cognitive reflection explains individuals' quality of inventory decision-making. Narayanan and Moritz (2015) find that the cognitive profile of decision makers contributes to the bullwhip effect in a beer distribution game. Multi-echelon supply chains managed by individuals with higher cognitive reflection have lower costs, exhibit less order variance and have lower demand amplification. B. B. Moritz et al. (2013) investigate the relationship between cognitive reflection and newsvendor decision making. They find that individuals with higher cognitive reflection exhibit a lower tendency to chase demand and that the cognitive reflection is a better predictor of performance. B. Moritz et al. (2014) also find that decision makers with higher cognitive reflection tend to have lower demand forecast errors. All being important and stochastic inventory environment where individuals' with lower cognitive ability make relatively poor decisions, we are interested in knowing whether different levels of thinking have an impact on the more static and non-strategic EOQ environment.

The EOQ model is developed by Harris (1913), it is simple yet commonly used in industries that manage durable goods, for example, furniture, household goods, lightning products, clothing and books. Our study provides a much broader managerial applications to practical problems as comparing to managing perishable goods. Many popular enterprise resource planning (ERP) software packages use the EOQ model as built-in calculation for planning and inventory controls. In Chapter 3, we reported on interviews with inventory managers from Chinese durable goods manufacturers, which confirmed that the ERP systems and the EOQ model are the predominant tools in the industry for inventory decision support. Our experimental results demonstrated that inexperienced inventory managers learn to achieve the expected profits and adopt the dynamic optimal strategy in the deterministic experimental setting. In the research paper by Pan, Shachat, and Wei (2019), further estimations suggest that participants are less likely to choose EOQ cycle lengths that increase payoffs under high cognitive load. The estimations also show that with the more complicated policy choice sets of the Unrestricted treatment, participants are more reluctant to make large changes in EOQ cycle length, and had a greater degree of strategy lock-in which reduces the probability of switching to better strategies.

In the current chapter, we investigate the relationship between the individual decision maker's cognitive reflection and their performance in inventory management decision making. As suggested by dual process theory Stanovich and West (2000), individual decision making often involves an interaction between two systems - System 1, identify with intuition, and System 2, identify with deliberation and reasoning. We measure subjects' cognitive ability by their Cognitive Reflection Test (CRT) scores. The CRT is a widely used performance-based measure designed to assess individual's tendency to override impulsive but wrong responses (System 1), and engage in a more effortful and reflective thinking for the right answers (System 2). The prediction of System 1 thinking in the context of inventory management task is to meet the monthly demand, while the prediction of System 2 thinking is to meet the monthly demand as well as minimising the total costs, consequently, maximising the profits. Frederick (2005) suggested CRT can be used as a simple measure of a person's cognitive ability and is correlated with academic performance such as SAT and ACT scores. Other studies have shown that individuals with higher CRT scores are less affected by heuristics (Toplak et al., 2011), behavioral biases (Hoppe & Kusterer, 2011; Oechssler et al., 2009), anchoring (Bergman et al., 2010), and prone to perform dominant strategy according to the Nash Equilibrium in the beauty contest game (Brañas-Garza et al., 2012). It has also been proved that CRT test is not equivalent to a pure numeracy test (Campitelli & Gerrans, 2014; Liberali et al., 2012). Also see Brañas-Garza et al. (2015) for a meta-study of 118 CRT studies.

Despite its widespread application in supply chain management, there is few existing literature that examine the practice of EOQ model. We have discussed previously, Stangl and Thonemann (2017) used the EOQ model as one of the three environments to evaluate how different performance metrics from supply chain management that include equivalent information can affect the performance of actual human inventory decision makings. Another more recent experimental study based on the EOQ framework examines inventory management learning behaviour under both stable and changing environment (K.-Y. Chen & Wu, 2017). In Chapter 3, we implemented a 2x2 experimental design that examined exogenous imposed cognitive load on participants' cognitive resource, and the effect of restricting participants to EOQ type of order behaviour where they can only place an order once inventories are depleted. In the current research, we adopted our original experimental design on the inventory task, while only changing alternative holding and inventory costs which change the optimal inventory policy. When evaluating how well subjects are performing in a decision task, we consider whether they are able to quickly react to the information given with an impulsive response or if they can override the impulsive but intuitive response and exert more effort to find the correct answer.

Our results show that participants with higher CRT scores achieve higher profits on average, while participants with lower CRT scores exhibit faster learning rate in realizing the expected profits across iterations of the inventory management task. The advantage of higher CRT score participants diminishes after several repetitions. There is a trend that participants learned to adopt EOQ optimal and near optimal policies with experience. The learning mostly occurred in the first three iterations of the task. Consistent with previous literature (Cueva et al., 2016; Primi et al., 2016), we found females have lower CRT scores on average, leading to an initial gender performance gap in the inventory task. It is worth mentioning that de Vericourt et al. (2013) have noted gender differences in newsvendor's inventory management decision making. They argued that female participants are greater risk averse which is a major driver of these differences. In the high margin settings of a newsvendor problem, risk taking behaviour is rewarded in payoff. Male participants tend to order more, thereby achieving higher profits.

In our deterministic EOQ settings, the solution is invariant to a decision maker's risk attitude and strategic considerations. Whereas it is not true in the newsvendor problem and the multi-echelon beer game as differences in risk attitude and strategic considerations are time invariant. In our experiment, we provide participants with simple and consistent feedback which eliminates the negative impact of demand chasing. Pure experience from practice and feedback can compensate for the disadvantage of hiring females, consequently, reduce gender under-representation. Our analyses proved that there is no gender effect within each CRT levels. Female participants improved more rapidly through consequential feedback and there is no gender difference in terms of profitability after a few repetitions of the task. In current workplaces, personality and behavioural types of cognitive test have become popular in job screening when companies are hiring people for roles, key roles tend to be assigned to applicants with higher scores. However, CRT types of tasks are correlated with gender, if one pursues such policy it would lead to gender biases in selection, which is not ideal for workplace diversity. This practise creates discrimination against the female group, as their performance will not differ from the male group on average after building a few round of experiences. Level of CRT scores matters but only at the beginning when participants first face the inventory management task. Participants with lower cognitive abilities actually learn and improve rapidly in terms of profits generated.

We conducted further interviews with inventory managers from three Chinese national firms operating in furniture, clothing and household hardware industries. The managers testified the existence of workplace gender inequality in supply chain management and the potential discrimination arising from current recruitment processes. We propose in this research study that organisations need to refine their selection processes for inventory management roles to prevent biased recruitments in the first place, rather than spending money to close the gender pay gap after the fact. It would create a biased selection if they put emphasis on testing cognitive measurements in terms of employment decisions. We suggest companies to be cautious of using cognitive ability tests when hiring for inventory management roles, and should take necessary training and probationary period into consideration.

4.2 Experimental design

The experiment consists of three tasks, a cognitive reflection test (CRT), an ego depletion task [9], and an inventory management task. The design of the inventory management task followed our original deterministic EOQ experimental settings. In the following discussions, subjects will be classified into three groups according to their performance in the CRT task. All experimental tasks were incentivised by monetary payoffs. The conversion rate for experiment currency unit and GBP is 450 ECU = £1 cash payment.

4.2.1 Cognitive Reflection Task

The first task is a standard cognitive reflection test (CRT). We followed the test developed by Frederick (2005). Subjects have three minutes to answer three short questions. Each correct answer gains 300 ECU. The CRT task measures a person's ability to overwrite an impulsive but wrong answer and engage in a more reflective thinking to find the correct answer. The correct answers are often found to be easy to understand once explained. Being successively resisted to give the intuitive spontaneous response is indicating that the individual is willing to engage in a more deliberate decision-making. The three questions are given as follows.

1. A bat and a ball cost 22 dollars in total. The bat costs 20 dollars more than the ball. How many dollars does the ball cost? (impulsive answer: 2; correct answer: 1)

2. If it takes 5 machines 5 minutes to make 5 widgets, how many minutes would it take 100 machines to make 100 widgets? (impulsive answer: 100; correct answer:5)

3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how many days would it take for the patch to cover half of the lake? (impulsive answer: 24; correct answer 27)

4.2.2 Inventory decision task

The core decision-making part in our experiment is an inventory management task, which follows the design in section 2.2, where participants were asked to make monthly ordering decisions for a consecutive of six years. We have used a different set of parameter values in the task:

1. The constant demand rate (D) of S-store is 20 units per month.

2. S-store sells coffee makers at a price of ₱5 per unit.

3. A fixed ordering cost (S) of ₱80 is incurred every time an order is placed, and does not depend upon the size of the order.

4. A variable monthly inventory holding cost is also paid based on the average number of coffee makers held in inventory multiplied by the per unit monthly inventory holding cost (h) of ₱0.5.

Therefore, the monthly profit of S-store is calculated as:

[9] We plan to use the data collected from this task in Chapter 5 of this thesis.

$$\pi_t(q_t, I_{t-1}) = \begin{cases} 5 \cdot 20 - S \cdot \mathbb{1}_{q_t>0} - \frac{I_{t-1}+q_t+I_t}{2} \cdot 0.5 & \text{if } I_{t-1} + q_t \geq 20 \\ 5 \cdot (I_{t-1} + q_t) - S \cdot \mathbb{1}_{q_t>0} - \frac{I_{t-1}+q_t}{2} \cdot 0.5 & \text{if } I_{t-1} + q_t < 20 \end{cases}$$

where, $\mathbb{1}$ is the indicator function.

A participant i's inventory policy for year a is the sequence of the twelve monthly quantity orders, $Q_{i,a} = (q_{i,1}, q_{i,2}, \ldots, q_{i,12})$. For a given inventory policy set S-store's annual profits is calculated as,

$$\Pi_{i,a}(Q_{i,a}) = \sum_{t=1}^{12} \pi_t.$$

Based on Equation 3, taking the parameters in this experiment, $n^* = 3$, $t^* = 4$, when the order decision is made for the whole finite horizon ($T = 12$). Thus, the optimal order quantity $Q^* = 80$. Within the finite horizon T, we derive that n^* is the smallest integer that satisfies

$$n(n+1) \geq \frac{hDT^2}{2S}$$

In our experiment, the optimal policy is EOQ constant 80, $Q^* = 80$.

There is a dynamic solution to the finite horizon EOQ Model conditional upon the starting month. If an inventory manager deviates from the EOQ policy early in the year then the optimal continuation course can involve alternative EOQ actions later in the year. Further investigations on situations when the horizon T is sufficiently small reveals that the optimal number of orders, n^*, is the smallest integer satisfying $n(n+1) \geq \frac{hDT^2}{2S}$. With the parameter values in Chapter 4, Table 9 gives an overview of the optimal solutions for different values of T.

Table 9: Optimal solutions for different T in our task

Month	T	$\frac{hDT^2}{2S}$	$n^*(n^*+1)$	The optimal number of orders (n^*)	The optimal EOQ cycle length (s_k^*) sequence	The optimal order size (q_k^*)
12	1	0.063	2	1	{1}	{20}
11	2	0.25	2	1	{2}	{40}
10	3	0.563	2	1	{3}	{60}
9	4	1	2	1	{4}	{80}
8	5	1.563	2	1	{5}	{100}
7	6	2.25	6	2	{4, 2}	{80, 40}
6	7	3.063	6	2	{4, 3}	{80, 60}
5	8	4	6	2	{4, 4}	{80, 80}
4	9	5.063	6	2	{4, 5}	{80, 100}
3	10	6.25	12	3	{4, 4, 2}	{80, 80, 40}
2	11	7.563	12	3	{4, 4, 3}	{80, 80, 60}
1	12	9	12	3	{4, 4, 4}	{80, 80, 80}

where, s_k^* is the optimal EOQ cycle length, and q_k^* is the optimal order size.

Table 10 compares the efficiency of those EOQ constant policies with no stockouts incurred in any month, or any positive closing inventories in storage by the end of month 12. The optimal inventory policy has 100% efficiency in terms of its annual profits generated. Notice that EOQ constant 120 and EOQ constant 60 can also generate over 94% of the optimal annual profits. Given the minimal loss incurred by adopting these policies, we define EOQ constant 120 and EOQ constant 60 as "near optimal" performance.

Table 10: Alternative EOQ policy performance efficiency table

EOQ constant	Orders per year	Annual profit	Efficiency
240	1	400	55.56%
120	2	680	94.44%
80	3	720	100.00%
60	4	700	97.22%
40	6	600	83.33%
20	12	180	25.00%

An important aspect of the inventory task is that we require the participants to spend a fixed amount of time in completing each inventory order decisions. The time restriction on each month in Years 1 through 5 is exactly 30 seconds. Participants can order any quantities every month that does not exceed the annual demand. If a participant spent less than 30 seconds in a particular month, then the participants would jump to a wait page and review the results of the current month. If a participant does not wish to place an order this month, then the participants simply do not enter an order quantity and wait until time expires, the computer program automatically executes the current month sales with the existing inventory in stock. The design requires more exertion of self-control when the participant wants to order zero quantity, consequently, to perform an EOQ policy strategy.

4.2.3 Experimental procedures

The experiment was conducted at the Experimental and Behavioural Economics Lab of Newcastle University. The participants were students from Newcastle University. Four experimental sessions were conducted in October 2017, and each session spent 90 minutes. In total, 113 subjects participated in the experiment. The subjects were recruited from the subject pool of the Behavioural Economics Northeast Cluster via invitation emails. Participants were not allowed to communicate with each other during the experiment, and they could not use pen and paper, calculator or a mobile phone. The experiment was programmed and conducted using a self-contained programme developed in oTree (D. L. Chen et al., 2016). They were further instructed that the experimental tasks are fully computerized. The information displayed on their computer monitor is private and the decisions are made individually. Access to other programmes on the computer was restricted. Participants can ask questions to an experimenter at any point during the experiment session. Once instructed to start by the experimenter, participants started to read through the instructions[10] at their own pace. After reading the instructions, participants started with the CRT task, followed by the ego depletion task.

[10]In the supplement appendix, we provide a complete set of instructions.

Participants then start with the inventory management task. After reading the background information, participants were asked to complete 7 multiple choice questions designed to ensure that they understand the calculation of S-store's costs and profits. Participants who missed more than 2 correct answers had to review the questions with one of the experimenters before proceeding to the inventory decision tasks. Year 0 is a practice round which used an alternative set of parameters[11] from those of Years 1 through 5, and the performance in Year 0 is not related to a participant's total earnings. The purpose of the practice year is to give the participants an opportunity to get familiar with the software and decision tasks. To help participants with understanding the task, we provide basic formulas, key information regarding the current month, and a monthly record table that calculates the sales revenue, costs, and profits on the decision screen, as well as an annual profit table that displays the profits made in each year from Years 1 through 5[12].

At the end of the experiment participants will be paid £5 show-up fee and their accumulated earnings from the three tasks, converted to pounds. Negative profit may occur if poor inventory order decisions are made. There was limited liabilities; to ensure that no one leaves the experiment with a payment less than £5 or affected by a large negative earnings made in a particular year, any negative profits made in a year will be treated as 0 earnings. The average earning in the experiment is £17.22 per participant, including the show-up fee. The payments were handed out to participants in sealed envelopes, so that the amounts received in the experiment were not revealed to other participants.

At the end of the experiment we collected demographic information in a short survey, including participants' age, gender, level of education, if they have had a course on supply chain management, and if they are a native English speaker. Participants' average age was 21. Of the 113 participants, 65% were female, 21% were postgraduate, 8% had taken a course on supply chain management and 73% were native English speaker. We have found that the performance in the inventory management task does not related to the demographics of the participants; and participants' current state of mind did not affect their performance on the task.

4.3 Hypotheses

The optimal inventory management policy in our EOQ setting is the solution to a finite horizon dynamic decision-marking problem. It involves the decision to correctly balance the number of order placed and the level of inventory kept in stock that minimize the total amount of ordering costs and inventory holding costs, consequently maximize the expected profit of the store. The research study by Stangl and Thonemann (2017) demonstrated that while the focal aspect of ordering and holding cost can be nudged through framing, order costs are more cognitively salient to individuals. We expect individuals with higher levels of cognitive reflection to process the task with less judgement bias driven by their initial focal perceptions. Therefore, having a higher probability of avoiding the intuitive but wrong policy of simply ordering the monthly demand. We also expect them to avoid EOQ inconsistent monthly decisions such as allowing irrational stockouts or placing orders when initial inventories are greater than monthly demand.

[11] In the practice year the order costs were ₱90 and the holding costs were ₱1.

[12] We provide screen captures of these interfaces in the supplement appendix.

Hypothesis 5. *Participants with higher cognitive reflection have higher average annual earnings.*

Optimal or near optimal policies have higher returns efficiencies. The optimal policy has 100% efficiency in terms of its annual profits generated, while near optimal policies can also generate over 94% of the maximum annual profits. We expect individuals with higher levels of cognitive reflection to avoid the impulsive wrong ordering strategies which require no or low cognitive effort, and engage in a more reflective thinking for to find the policies that return higher profits.

Hypothesis 6. *Among the participants with higher cognitive reflection, the percentage of who adopt optimal (near-optimal) inventories is higher.*

4.4 Results

4.4.1 CRT task results

The CRT score is measured in terms of the average number of correct responses given by participant. According to Frederick (2005), individuals who had low CRT scores are prone to answer with the obvious answer that immediately comes to mind. The average CRT score in our sample is 1.11. Among our 113 subjects, 14.16% answered all three questions correctly, 21.24% answered two questions correctly, 25.66% answered one question correctly, and the remaining 38.94% gave no correct answers. Table 11 summarises the percentage of participants provided the impulsive but incorrect answer, comparing with the percentage of participants who provided correct answer for each question, respectively. Note that the impulsive answers were more popular than correct answers in the first two questions, and most of the participants gave either the correct or the impulsive answers for the questions.

Table 11: Percentage of correct and impulsive answers on the CRT test

Question	Correct	Impulsive	Other
Bat and ball	34.51%	60.18%	5.31%
Widgets	30.09%	49.56%	20.35%
Lily pads	46.02%	35.40%	18.58%

4.4.2 Results on inventory management tasks

We construct three categories of CRT performance for individual participants. Answering zero questions correctly is a "Level 1" performance, where participants gave the wrong initiative answers to all three questions. As opposed to "Level 1", "Level 2" performance includes participants engaged in further reflection to work out all three correct answers. The remaining participants who answered either one or two questions correctly are in the "Mixed" performance group, as they exhibited a mixed behaviour in reflective thinking. We report the average annual profits made by subjects by CRT scores in Table 12 Panel A. Participants with higher CRT scores achieved better performance in the inventory management task. The standard deviation in average profits is larger in the CRT-Level 1 group than the other two groups.

We test the differences in average annual profit for different treatment groups using two sided t-tests and non-parametric Wilcoxon rank-sum tests. The test results are reported in Table 12 Panel B. We find that participants' CRT score is positively correlated to their average annual profits. Two sided t-tests and Wilcoxon rank-sum tests showed evidence that such increase in average annual profit with CRT scores are statistically and economically significant.

Result 6. *Participants with higher CRT scores earn statistically significantly more.*

Figure 8 presents the average profits made in Year 1 through Year 5 by CRT score groups. We can observe learning over time and how does it different across groups. This figure has several prominent features that give us insight into the hypotheses results on average profits levels. The figure shows that at the beginning of the inventory management task, in Year 1, the average profits made by CRT-Level 1 group is around 20% less than the profits made by CRT-Mixed group, and around 30% less than the profits made by CRT-Level 2 group. The

Table 12: Average annual profits by CRT scores and hypotheses tests for differences in average annual earnings

Panel A: Annual profits by CRT scores

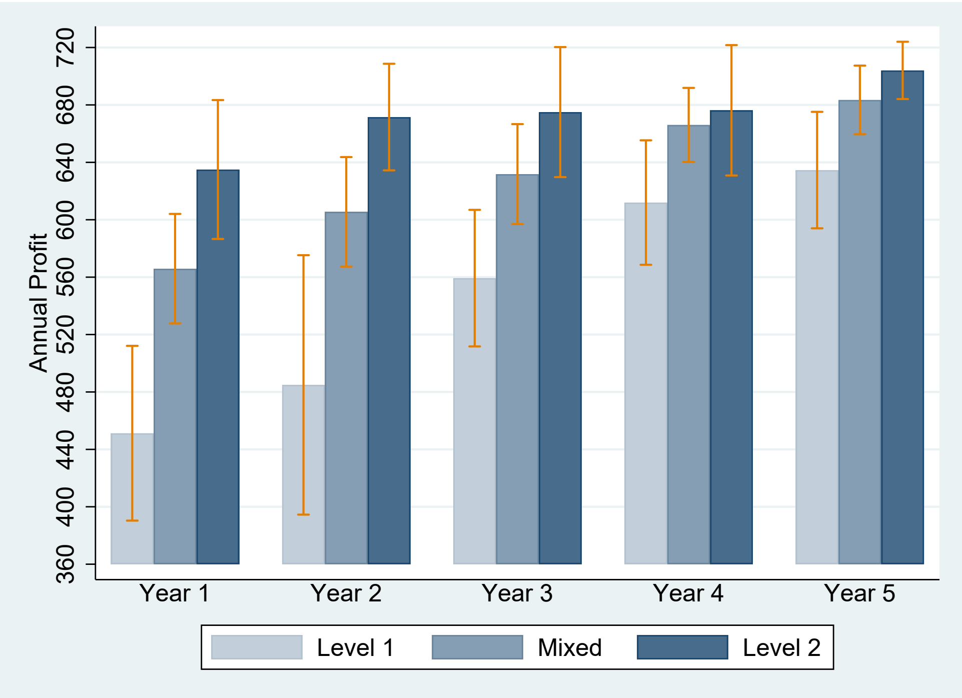
	Level 1	Mixed	Level 2
Average	548.41	630.56	672.37
Standard deviation	206.29	125.21	77.43

Panel B: Hypotheses tests for differences in average annual profits (p-values reported)

CRT Scores Comparison	Profit Difference	Two-sided t-tests	Wilcoxon rank-sum
Level 1 vs. Mixed	-82.14 (-14.98%)	0.000	0.000
Level 1 vs. Level 2	-123.96 (-22.60%)	0.000	0.000
Mixed vs. Level 2	-41.81 (-6.63%)	0.000	0.001

learning rate of the CRT-Level 1 group is the highest among the three groups. By Year 5, the average profits made by CRT-Level 1 group is only around 7% less than the profits made by CRT-Mixed group, and around 10% less than the profits made by CRT-Level 2 group. Subjects in the CRT-Level 2 group achieved high initial annual profits, around 88% of the optimal profit in Year 1. The average profits of the three groups tend to converge over time. The figure also suggests that learning is mostly occurred in Year 1 through 4, there is only marginal increase in average profits in Year 5. Also, the CRT-Level 2 group achieved higher than 90% of the optimal profits in Year 2; and the CRT-Mixed group achieved higher than 90% of the optimal profits in Year 4.

Figure 8: Annual Profits over individual Years and by CRT scores: Averages and 95% confidence intervals



We quantify our results by conducting a series of dummy variable linear regressions using random effects estimators and cluster standard errors at the level of the individuals. The regression results are shown in Table 13. In model (1) we regress annual profits on Year where the profit in Year 1 is included in the constant term. In model (2) we introduce dummy variables for the CRT group variables. The constant reflects the average annual profits of the CRT-Level 2 group in Year 1; the CRT-Level 1 and CRT-Mixed dummy variables reflect the average annual profit of the two groups in Year 1. In model (3) we include the interaction terms of the Year variable with the CRT group dummy variables for the additional learning rate of the CRT-Level 1 group and the CRT-Mixed group.

From the regression results we observe significant performance improvement on average each year. We can conclude that subjects who achieved higher CRT scores made higher annual profits on average, thus support Hypothesis 5. There is a significant tendency of learning from Year 1 through Year 5, and the subjects in the CRT-Level 1 group has the fastest learning rate. Further, we conduct a Chow test for the significance of the difference between model (2) and (3). The result supports the differences in annual profits across CRT groups (χ^2 (4, N=565) = 32.09, $p = 0.000$).

Table 13: CRT gorup dummy variable regressions for annual profit. (n=565)

	(1) Annual Profit	(2) Annual Profit	(3) Annual Profit
Year	35.12***	35.12***	14.28***
	(3.54)	(3.55)	(4.23)
CRT-Level1		-123.96***	-194.15***
		(25.55)	(36.09)
CRT-Mixed		-41.81***	-72.39***
		(15.87)	(22.09)
CRT-Level1*Year			35.10***
			(8.27)
CRT-Mixed*Year			15.29***
			(5.71)
Constant	534.25***	602.13***	643.81***
	(16.28)	(13.23)	(15.46)
chi2	98.19	102.66	125.62
p	0.00	0.00	0.00

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

4.4.3 Gender Effect

We have found consistent results with Frederick (2005) when comparing the results between genders. Male participants exhibit more cognitive reflection than the females. Table Table 14 reports the mean and empirical distribution of correct CRT question responses overall and then for the female and male participants. The mean CRT score for males is 1.51, which is significantly higher than that of females at 0.89 (Wilcoxon rank-sum test, p = 0.003). Our categorization also highlights gender differences in latent cognitive reflection. There are around 49% of the female participants in the CRT-Level 1 group, whereas that is only around 21% for male participants. In contrast, the proportions of female and male participants in CRT-Level 2 group are 11% and 21% respectively. We will consider this correlation between gender and CRT performance in later discussions when examining inventory management performance and behaviour.

Table 14: CRT Scores

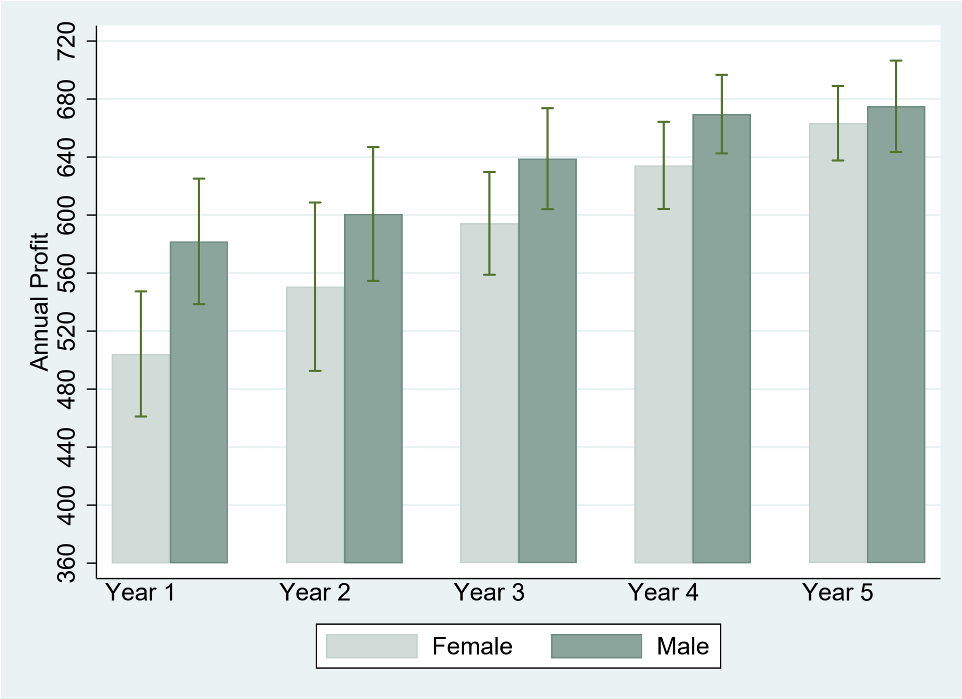
		Percentage scoring 0, 1, 2, or 3				
		"Level 1"	"Mixed"		"Level 2"	
	Mean CRT score	0	1	2	3	$N =$
Overall	1.11	39%	26%	21%	14%	113
Female	0.89	49%	24%	16%	11%	74
Male	1.51	21%	28%	31%	21%	39

To examine whether there is a gender effect in the inventory management task, first we consider the differences in average annual profit between Male and Female participants. The average annual profits of Female and Male participants are 589.35 and 633.23 respectively. This difference of 43.88 is statistically significant according to both a t-test ($p = 0.001$), and a Wilcoxon rank-sum test ($p = 0.027$). However, Figure 9 suggests that these differences are more pronounced in earlier years. In fact, we conducted both t-tests and Wilcoxon rank-sum tests for gender differences in each year. We only reject at the 5% level of significance that there are no differences in annual profit between male and female groups in Year 1.

Next, to quantify our result we introduce the gender dummy in model (1). The constant reflects the average annual profits made by male participants. In model (2) we include dummy variables for the CRT group variables. The constant reflects the average annual profits of the CRT-Level 2 group; the CRT-Level 1 and CRT-Mixed dummy variables reflect the average annual profit of the two groups. In model (3) we include the interaction terms of the gender variable with the CRT group dummy variables to examine the gender difference in average annual profits within each CRT groups.

From the regression results in Table 15 we can conclude that there is no gender effect in the inventory management task. As suggested by model (1), female participants achieved lower annual profits on average. However, the Female coefficient becomes insignificant when we include CRT level variables in model (3). This suggests that there is omitted variable bias in model (1) and the drive of differences in profit is CRT level rather than gender. Male participants made higher profits on average than female participants due to a large portion of female participants

Figure 9: Annual Profits over individual Years and by gender: Averages and 95% confidence intervals



are in the CRT-level 1 group; male participants have higher CRT scores than female participants on average. However we cannot observe gender differences in average annual profit within each CRT groups. We proved using a Chow test for the significance of difference between model (1) and model (3) (χ^2 (4, N=565) = 24.04, p = 0.000). Further, we conduct a second Chow test for model (2) against model (3). The result confirms that introducing the interaction variables is not justified which supports no gender differences in average annual profit (χ^2 (3, N=565) = 1.10, p = 0.778).

Table 15: Gender dummy variable regressions for annual profit. (n=565)

	(1) Annual Profit	(2) Annual Profit	(3) Annual Profit
Female	-43.88**		-9.99
	(20.87)		(21.59)
CRT-Level1		-123.96***	-110.15**
		(25.53)	(52.98)
CRT-Mixed		-41.81***	-36.54*
		(15.85)	(19.11)
Female*CRT-Level1			-12.99
			(60.74)
Female*CRT-Mixed			-8.16
			(30.42)
Constant	633.23***	672.37***	677.37***
	(13.74)	(10.84)	(16.06)
chi2	4.42	25.23	25.38
p	0.04	0.00	0.00

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

The performance of inventory management significantly differs, in terms of annual profits made, when comparing between male and female groups across all CRT categories. However, the differences are driven by differing performance in earlier years. As years progress Level 1 and Mixed CRT participants' performances improve more rapidly; performances tend to converge to the same level by Year 5. We also provide evidence that the gender performance gap is actually misleading. We have modelled the performance between gender groups using multivariate regressions. The correlation between gender and performance reflects an omitted variable problem rather than a true correlation. This leads to an additional problem with using CRT types of questions in job screening, it excludes more females when in fact their performances will not differ from men on average after gaining simple experiences.

4.4.4 Inventory management policy choices

CRT score measures the tendency of avoiding an impulsive but wrong answer and engage in a more reflective thinking to find the correct answer. In our experiment, an intuitive inventory policy would be to take the opportunity to order in each month, because an impulsive decision is to order the monthly demand not taking consideration of other costs that may generate; thus, to order 20 units every month. Our analysis also shows that the percentage of CRT-Level 1 participants that adopt the EOQ constant policy is higher than other CRT groups. However, we didn't observe such difference with other EOQ policies. We analysed participants' inventory policy choices and found that the percentage of choosing EOQ constant policy 20 is 11.36% in CRT-Level 1 group; while the percentage in CRT-Mixed and CRT-Level 2 groups are 7.55%, and 0%. Participants with higher CRT scores successfully avoided the impulsive, therefore, less efficient inventory ordering policy.

We have analysed each participant's inventory policy choices on an annual basis. For whether $Q_{i,a}$ is optimal, if $Q^* = 80$; or $Q_{i,a}$ is near optimal, if the annual order decisions are any combinations of EOQ policies of 60, 80 and 120. Figure 10 shows the percentage of participants following optimal and near optimal strategies in each CRT score groups. In general, the percentage of participants adopting optimal or near optimal policy is higher in the CRT-Level 2 group than in the CRT-Mixed group and the CRT-Level 1 group. We can see a clear trend in the increase of adopting optimal or near optimal policy from Year 1 through Year 5. The percentage of participants who adopted the optimal policy is approximately the same in Year 1 for three different groups; while the CRT-Level 2 group has a faster learning rate. For participants in the CRT-Mixed and CRT-Level 2 group, the increasing trend of adopting optimal or near optimal policy occurred in the first three years of the inventory task. For participants in the CRT-Level 1 group, the increasing trend of adopting optimal or near optimal policy occurred in the last three years of the inventory task.

Different from the experimental design in Chapter 3 where we eliminated Non-EOQ decisions in the Zero-Only treatments, participants in this experiment need to actively make decisions under the EOQ policy. We turn our analysis to how different CRT levels impact individual learning process. We investigate whether difference in CRT scores is influencing the probability of subjects to adopt EOQ optimal or near optimal policy. The Logit regression results are presented in Table 16.

First, as indicated in the last row of the table, the average levels of the estimated probability of adopting an optimal policy are lower than the probability of adopting a near optimal polity. Second, the coefficients on Years are positive and significant indicating there is a significant trend in learning to make EOQ decisions from Years 1 through 5. Third, the coefficients on CRT-Level 1 and CRT-Mixed are negative and significant, this is indicating that the probabilities of adopting optimal or near optimal policies are the highest in the CRT-Level 2 group, lower in the CRT-Mixed group and lowest in the CRT-Level 1 group, thus confirming Hypothesis 6. Lastly, in model (3) and model (6) we include the gender dummy, and conclude that the gender effect is insignificant across all models.

Figure 10: Stacked graph of the percentage of participants following optimal and near-optimal EOQ constant policies: by Year and CRT score

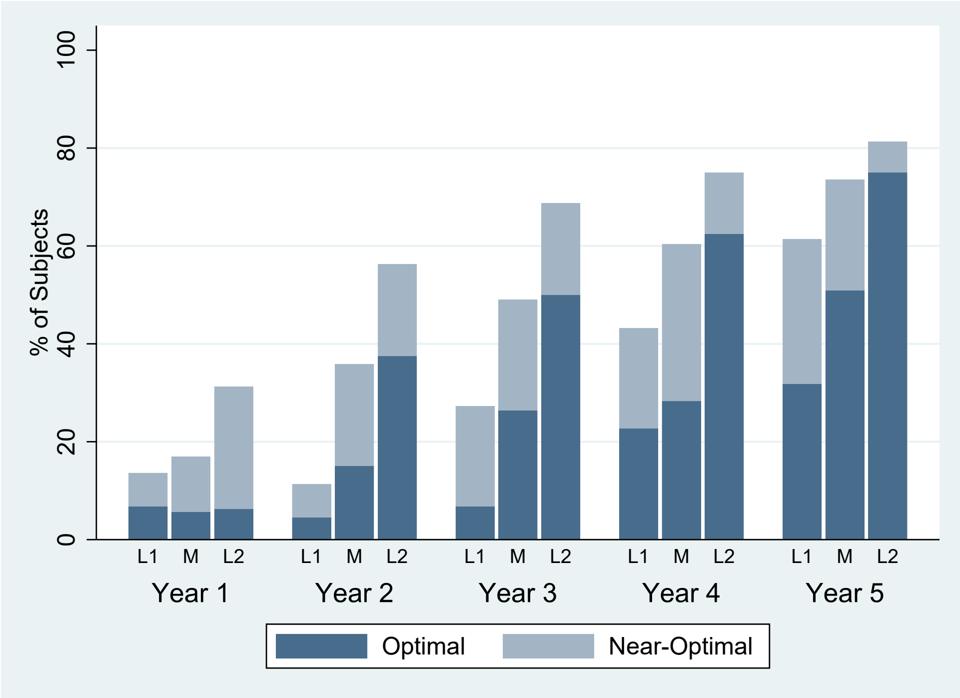


Table 16: Logit regression on the probability of choosing optimal or near optimal policy. (n=565)

	(1) Optimal	(2) Optimal	(3) Optimal	(4) Near-optimal	(5) Near-optimal	(6) Near-optimal
$Year$	0.597***	0.643***	0.644***	0.584***	0.621***	0.621***
	(0.062)	(0.069)	(0.069)	(0.061)	(0.067)	(0.067)
CRT-Level1		-1.868***	-1.928***		-1.526***	-1.510***
		(0.458)	(0.455)		(0.447)	(0.447)
CRT-Mixed		-1.099***	-1.113***		-0.739*	-0.736*
		(0.394)	(0.396)		(0.417)	(0.417)
Female			0.184			-0.0509
			(0.317)			(0.287)
Constant	-3.117***	-2.109***	-2.204***	-2.072***	-1.258***	-1.232***
	(0.269)	(0.373)	(0.401)	(0.236)	(0.382)	(0.406)
χ^2	92.72***	89.73***	92.33***	91.30***	86.47***	86.77***
Pr	0.241	0.241	0.241	0.432	0.432	0.432

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Result 7. *There is a trend for all three CRT score groups for increasing use of optimal and near-optimal policies from Year 1 to Year 5.*

Result 8. *Participants have higher CRT scores tend to have higher percentage use of optimal and near-optimal policies in all five years.*

Finally, we discuss how participants' CRT levels impact their Non-EOQ decisions. We followed the decision process for monthly choices in the analysis of Chapter 3. Under our current experimental setting, the decisions is made upon whether the closing inventory of previous month is greater or less than 20, where the Non-EOQ decisions result from either stockouts[13] or holding excess inventory.

We estimate a set of Logit regressions on the probability of a subject makes Non-EOQ decision. The probabilities of making Non-EOQ decisions are formulated as simple Logit functions of a function of time, accumulated Non-EOQ ordering habit, and different CRT groups. In model (3) and (6) we introduce variables to capture the habit of making Non-EOQ decisions. $NonEOQACC_{i,r-1}$ is the running count of the total number of rounds the subject i has deviated from the EOQ decision up to the round $r-1$, which intend to capture the habit formation. The Logit regression results are presented in Table 17: Panel A for the case $I_{t-1} < 20$ where there are possibilities of stockouts if $q_t < 20 - I_{t-1}$, and Panel B for the case $I_{t-1} \geq 20$ where $q_t > 0$, leading to excess inventory.

First, from the regression results we have large negative values of the estimated constant, thus, the average levels of the estimated probability of making a Non-EOQ decision are very small as indicated in the last row of the table. Second, the coefficients on Years are negative and significant indicating there is a significant trend in learning to make EOQ decisions from Year 1 through Year 5. The estimated coefficients on Months are significant with smaller magnitudes, and have opposite signs under the two conditions. This suggests that stockouts are more likely to happen later in a year while having excess inventory in stock is less likely to happen later in a year. Third, the estimated coefficients on the accumulated Non-EOQ decisions are positive and significant captures the individual differences in adopting the EOQ ordering logic. Lastly, we turn our analysis to whether participants with lower CRT scores being more likely to make Non-EOQ monthly choices. We see the estimated probabilities of these two errors are higher for CRT-Level 1 and CRT-Mixed groups. However, this is only statistically significant for the case of incurring unnecessary holding cost by ordering when initial inventory exceed 20.

Result 9. *CRT Level 1 participants are more likely to make Non-EOQ decisions when initial inventories are greater than monthly demand.*

4.5 Managerial Application

We have conducted interviews with three national firms in the furniture, clothing and household hardware industries based in Guangdong Province, China. The inventory managers use ERP system as the predominant tool to assist their management of inventories. They rely on the system's records to keep track of the inventories in stock, sales and ordering activities. However, they tend to coordinate with the manufacturers and make actual ordering decisions to overwrite the deterministic solution from the system. In the case of introducing new systems or new products launch, managers need to carry daily reviews on the system with the assistance of manual accounting. After the first year trail, managers stop doing manual accounting and rely entirely on the system. They switch from daily reviews to monthly reviews; account balancing

[13] We recognise that with our setting, in later months, it may be more profitable to suffer a stockout when the open inventory is not too far short from the demand. For example, if a participant's opening inventory is above 15 units but less than 20 in month 9, it would be more profitable to suffer a stockout and wait until month 10 to order 60 than order the amount short from 80. This may lead to a situation where participants deliberately suffer a stockout. However, out of 6780 observations, only one observation matches the situation.

Table 17: Logit regression on the probability of deviating from an EOQ action

	Panel A: $I_{t-1} < 20$			Panel B: $I_{t-1} \geq 20$		
$NonEOQ_{i,r}$	(1)	(2)	(3)	(4)	(5)	(6)
$Year_r$	-0.176**	-0.166**	-0.283***	-0.484***	-0.512***	-0.990***
	(0.081)	(0.081)	(0.101)	(0.061)	(0.063)	(0.095)
$Month_r$	0.152***	0.148***	0.142***	-0.0844***	-0.0878***	-0.144***
	(0.038)	(0.038)	(0.042)	(0.020)	(0.020)	(0.024)
CRT Level 1		0.646	0.315		2.138***	1.517***
		(0.445)	(0.456)		(0.394)	(0.408)
CRT Mixed		0.0337	-0.153		1.093***	0.983**
		(0.470)	(0.478)		(0.402)	(0.406)
$NonEOQACC_{i,r-1}$			0.113***			0.279***
			(0.029)			(0.021)
Constant	-3.818***	-4.166***	-3.828***	-1.291***	-2.718***	-1.595***
	(0.400)	(0.541)	(0.597)	(0.207)	(0.414)	(0.416)
N	2008	2008	1895	4772	4772	4772
χ^2	26.65***	32.88***	51.84***	75.67***	130.8***	241.2***
$Pr(NonEOQ_{i,r}) = 1$	0.0344	0.0344	0.0348	0.0417	0.0417	0.0417

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

within the predetermined error margin can be accepted.

As we have raised the concern of workplace gender inequality in supply chain management, the interviewees also testified that over 90% of the inventory managers in their companies are male managers. This number is consistent with the 2004 annual survey of logistics managers and directors in the US as discussed by Mangan and Christopher (2005). There are many historical or naturally existing reasons behind workplace female under-representation, for example the education effect, the occupation effect, the part-time work effect, and the motherhood effect. However, the potential discrimination existing in the recruitment processes has raised little concern. Finally, we have asked the managers to rank the top three employment attributes that companies are looking when hiring for inventory managers; adaptability and flexibility was ranked the top, followed by reasoning and problem solving skills, and stress management skills. The top attributes for inventory managers are in fact measured by cognitive ability tests, however, are not accessed based upon long-term performance. If we blindly follow this it would lead to disproportionally selection of female.

4.6 Conclusion

In Chapter 4, we studied the relationship between individuals' level of cognitive reflection and their performance in the EOQ inventory management task. We have provided statistical evidence to explain the gender gap in terms of the profits earned between male and female groups. The design of our experiment adopted the finite horizon EOQ setting which does not include risk, uncertainty or strategic issues. These permit evaluations of these relationships without considering the effects of individual risk attitudes, the resolution of strategic uncertainty, and the coordination among others.

Our results show that participants with higher CRT scores can earn more profits in the proceeding inventory management task, and can make more optimal ordering decisions. Male participants outperform female ones in these tasks. Beyond the initial observations on the performance gap, we noted that the differences of profits earned and adoption of optimal policy are not persistent. The gender gap closes through learning across several repetitions of our inventory management tasks. Consistent with previous studies, female participants exhibit lower CRT performance than males in our experiment. In a multivariate analysis that control for both gender and CRT performance, we find that differences in CRT performance are the main driver of earnings differences. Our finding has raised the concern of using cognitive ability tasks in the process of employee selection. In the absence of such measures one might mistakenly conclude that the quality of supply chain management performance is gender driven.

The current chapter has been initially motivated by a talk from a Harvard scholar, Dr Iris Bohnet, about 'workplace gender inequality' (Morse, 2016). The practical implication of our results provides managerial guidance to firms and organisations that aims to improve female under-representation and perhaps other under-represented groups, in particular supply chain management roles. We propose in this research that organisations need to refine their selection processes to prevent biased recruitments in the first place, rather than spending a large amount of money in closing the gender pay gap after the fact. Firstly, one should be aware of not to underestimate individuals' learning ability, and the improvement they can achieve from accumulated experiences. Second, one should be cautions of those recruiting practices that eliminate individuals from certain groups, as such practices lead to selection biases and workplace inequality.

There are some limitations to the current designed of the experiment. As discussed in the result section, performance gaps between different CRT level groups tend to close through Year 1 to 5, the differences are only marginal. However, the gap between CRT-Level 1 and CRT-Level 2 people are still visible in year 5. Moreover, although the difference between percentage of subject adopting near optimal policies decreased across years, there is still a significant difference in terms of percentage of subject adopting optimal policy. A potential future research project would be to study to what extend would a longer period of training close the performance gap between different CRT groups as subjects accumulate experience over time. Our current research suggests that although cognitive screening works effectively in recruiting processes, however, we should be aware of the potential trade-off of adopting such methods in recruitment. For enterprises that a recruiting for inventory manager type of roles, CRT type of cognitive tasks have the potential of resulting in gender biased selections. The managerial objectives of inventory managers are concentrating on a long-term success, rather than a short-term instance response. A reasonable next step is to find out what happens if more years are played in the experiment, therefore, provide evidence of how longer terms of training could mitigate the performance gaps.

In the next chapter we are going to provide a discussion on the results from the ego depletion treatment.

5 Ego Depletion

5.1 Introduction

We provide a discussion about the ego depletion treatment we have implemented in a separate chapter. In Chapter 5 we examine self-control and the impact of ego depletion upon decision quality in inventory management. Ego depletion comes into play at work when inventory managers become experienced and well learned in giving dominant responses in day-to-day decision making tasks. We predict that exercising self-control at work diminishes managers' ability to perform well on proceeding tasks.

When evaluating how well subjects are performing in a decision task the second factor we consider is once the optimal strategy is found, do subjects keep performing the strategy or they choose to deviate from the strategy. In the research study by Magnani et al. (2016) where there is a stochastic process of the inventory problem, the evidence showed that experienced participants found the optimal solution in a more dynamic (S, s) model, and they would actually switch away from the optimal. However, (S, s) model has limited explanatory power in explaining frequent errors and deviations occurred in dynamic decisions. In the condition where the feedback is stochastic, people may end up chasing small probabilities, random demand doesn't lead to consistent feedback. The research suggests that cognitive costs are much larger in higher volatility conditions, the volatility significantly increased the difficulty of the decision task, thus make it even costly to optimize. Furthermore, there is a conjecture about self-control issues arising from individuals' bounded rationality. From a psychological point of view, self-control is a limited amount of metal resource that is exhaustible. Self-control is impaired when the metal resource has been used up over effortful control of responses, then a person would be considered to be at a state of ego depletion that would affect his ability to control himself on subsequent tasks. In our experiment, the simplicity and deterministic nature of the EOQ framework provides inexperienced participants with the best chance to learn and adopt the optimal strategy.

Before the proceeding to the inventory management task, subjects were involved a variation of the letter 'e' task. The letter 'e' task was first introduced by Baumeister et al. (1998), and has been applied in charitable behaviour (Fennis et al., 2008), interpersonal relations (Tyler, 2008), persuasion resistance (Wheeler et al., 2007), the control of adverse emotions and actions (DeWall et al., 2007; Muraven, 2008). Also see Hagger et al. (2010) for a meta-study of 83 ego depletion studies. The treatment variable we implement is the level of ego depletion participants experienced in the task. In our Low-Depletion treatment participants follow a simple rule to delete the occurrence of the letter 'e' repeatedly throughout the task. In contrast, participants in Med-Depletion and High-Depletion face a more complex rule only to delete the word under certain conditions in stage two of the task. After having developed the habit of applying the simpler rule in stage one, overriding the previous habit involves more exertion of self-control. The ego depletion task corresponds to the condition when inventory managers become experienced and well learned in giving dominant responses in day-to-day decision making tasks.

In our experimental study, we examine for the first time the impact of ego depletion on decision maker performance on inventory management task. We frame a dynamic problem that has steady optimal controls in a deterministic EOQ environment. When doing repeating tasks it requires disciplines, following rules that need self-control, thinking slowly about the problem is more productive than reacting fast. We anticipated that the ego depletion task matters ex-ante, however, we do not observe performance differences or policy choice differences across ego depletion treatment groups. Subjects do not switch away from the better performing decisions once found, and there is no evidence of self-control issue in our static experimental setting.

5.2 Experiment

5.2.1 Ego-depletion task

The experiment follows a 3x1 experimental design. The treatment variable we implemented in the experiment is the level of ego depletion task. We adopted a between-subject design where a participant randomly takes part in one of the three treatments. The experimental task was incentivised by monetary payoffs. The conversion rate for experiment currency unit and GBP is 450 ECU = £1 cash payment.

We developed an ego-depletion task based on the letter 'e' task in the self-regulation research (Baumeister et al., 1998). The task consists of two parts. In the first part, subjects were instructed to delete words with the occurrences of the letter 'e'; otherwise keep the word. Part 1 involving 50 words with 10 words displayed on each page, subjects have up to 50 seconds to complete each page. This part is relatively easy and does not require much self-control. Subjects established the behaviour pattern to scan and delete every appearances of letter 'e'.

In the second part of the task, we implemented experimental manipulations to three different treatment groups. Part 2 involving 100 words and follows the same structure as in Part 1. Participants in the Low-depletion condition continued to do the task in Part 1 by applying the same rule that they had already familiarized with. Participants in the Med-depletion condition were instructed to delete words with the occurrences of the letter 'e', but keep the word if the letter 'e' occurs next to another vowel (e.g., read) or one letter away from another vowel (e.g. towel). This rule requires subjects to override the established behavioural pattern, and exert self-control to adapt to the new rule. We further developed a high-depletion condition as the third treatment group. Participants in the High-depletion condition were instructed to delete words with the occurrences of the letter 'e'. But keep the word if the letter 'e' occurs next to another vowel (e.g., read) or one letter away from another vowel (e.g. towel) or if it is also an adjective (e.g. excellent). This is a more complicated rule comparing to the Med-depletion condition that requires more self-control resource to override the existing habit.

Unlike previous self-regulation research, the major innovation in our design of the ego depletion task is that the task is incentivised, where subjects are paid upon the accuracy of the completion. Participants tend not to deliberately make random selections that require no self-control. We improve the traditional pen and paper approach by displaying the task on the computer screen. This way we are able to monitor and incentivise the accuracy by paying subjects the amount proportional to their performance. Participants tend not to make deliberate random selections that require no self-control. Each correct answer gains 15 ECU.

There was a short questionnaire after the ego depletion task, participants were asked to rate how hard it was to complete the task and how much effort did they put into completing the task on a 1-5 scale (1-not at all; 5-very much). Our experiment session took place in the afternoon. The questionnaire also concerns participants' current state of mind by asking how many hours did they sleep last night, what time did they wake up today, did they have breakfast, how many hours of lectures did they attend before coming to the experiment, and when did they eat their last meal.

5.3 Hypotheses

Participants were observed to make inconsistent decisions in a stochastic process of the inventory management problem Khaw et al. (2017); Magnani et al. (2016). They learn the optimal solution in a more dynamic (S, s) model with experience, however, their choices do not persist that they often subsequently deviate from the optimal policy. In our experimental setting, the simplicity and deterministic nature of the EOQ framework provides inexperienced participants with the best chance to learn and adopt the optimal strategy, which gives us the best condition to assess self-control issues correspond to individuals' bounded rationality in a static experimental setting.

Self-control is a limited amount of metal resource that is exhaustible over effortful control of responses. Exercising self-control in the ego depletion task reduces subjects' ability to perform well on the proceeding inventory management task. In our experiment, participants should engage in the ego-depletion task which requires effortful self-control since the task is monetary incentivised. We expect individuals who participated in lower level of ego depletion tasks to process the inventory management task with more self-control, therefore, they should hold until the current inventory equals to zero to place the next order, as opposed to making deliberate decisions in order to finish the rounds early. Consequently, achieving more profits by making EOQ consistent monthly decisions.

Hypothesis 7. *Participants who have engaged in lower level of ego depletion task will achieve higher average annual profits.*

Also, we expect individuals who participated in lower level of ego depletion tasks have more metal resource for the subsequent inventory management task. So that they would have better chance of finding optimal or near optimal ordering policies that have higher return efficiencies.

Hypothesis 8. *Among the participants who have engaged in lower level of ego depletion task, the percentage of who adopt optimal (near-optimal) inventories is higher.*

5.4 Results

5.4.1 Ego depletion task results

In the ego depletion task, the average percentages of the number of words correctly deleted or kept out of 150 words in three treatment groups are presented in Table 18. The differences between Low-depletion and the other two groups are both significant under two sided t-tests ($p = 0.023$ for Low vs. Med; $p = 0.001$ for Low vs. High) and non-parametric Wilcoxon rank-sum tests ($p = 0.000$); but no significance in difference was found between Med-depletion and High-depletion groups (t-test, $p = 0.210$; Wilcoxon rank-sum test, $p = 0.051$).

We implemented a manipulation check on the self-assessed questions followed by the ego depletion task. In terms of the perceived difficulty of the ego depletion task, the result from one-way analysis of variance (ANOVA) indicated significant variation among the three treatment groups, ($F(2, 110) = 26.97, p = 0.000$). The second question asked how much effort the participants felt they had put into completing the task. The differences in effort spent between groups were also statistically significant, ($F(2, 110) = 7.05, p = 0.001$). Thus, indicates our ego depletion manipulation was successful. The means and standard deviations are presented in Table 18. A Tukey post-hoc test revealed that self-reported difficulties were statistically different between treatment groups ($p = 0.000$). Note that subjects in the Low-depletion group reported less effort spent in completing the task comparing to the other two groups ($p = 0.000$); however, there were no statistical differences between Med-depletion and High-depletion group ($p = 0.145$).

Table 18: Ego depletion task completion accuracy and self-assessed quetions (scale of 1 to 5)

Treatment Groups	Average Accuracy	Difficulty	Std. Dev	Effort	Std. Dev
Low Depletion	97%	1.47	0.76	2.74	1.11
Medium Depletion	94%	2.21	0.84	3.63	1.22
High Depletion	92%	2.86	0.86	3.57	1.14

5.4.2 Annual inventory profits earning dynamics

We report the average annual profits made by subjects by ego-depletion treatment groups in Table 19 Panel A. The average inventory profits are on the same level, and the subjects in the High-depletion group had higher standard deviations in annual profits than the other two groups.

We test the differences in average annual profit for different treatment groups using two sided t-tests and non-parametric Wilcoxon rank-sum tests. The test results are reported in Table 19 Panel B. We find that the Med-depletion group makes higher annual profits in average than both Low-depletion and High-depletion group but the differences between groups are not significant. Implementing the ego-depletion task did not have a significant negative impact on average annual profits.

Result 10. *Participants' performance in the inventory management task did not differ in terms of the level of ego depletion task they have engaged in. Participants who have engaged in lower*

Table 19: Average annual profits by treatment and hypotheses tests for differences in average annual earnings

Panel A: Annual profits by treatment

	Low Depletion	Medium Depletion	High Depletion
Average	600.28	615.72	597.29
Std. Dev.	151.10	145.97	192.15

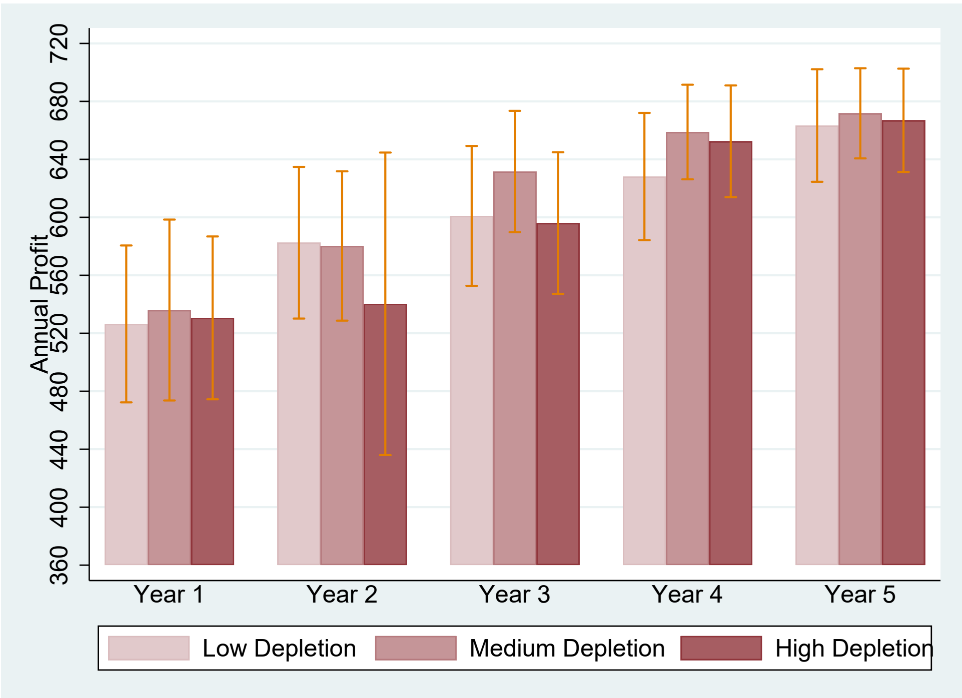
Panel B: Hypotheses tests for differences in average annual profits (p-values reported)

Treatment Comparison	Profit Difference	Two-sided t-tests	Wilcoxon rank-sum
Low vs. Medium Depletion	-15.43 (-2.57%)	0.312	0.041
Low vs. High Depletion	3.00 (0.50%)	0.867	0.537
Medium vs. High Depletion	18.43 (2.99%)	0.297	0.174

level of ego depletion task did not achieve higher average annual profits.

Figure 11 presents the average profits made in Year 1 through Year 5 by ego depletion treatment groups. A disaggregated view of the average annual profits permits insights into learning over time and how our treatments impact it. We can observe learning occurred from Year 1 through Year 5, whereas the average annual profits made by different ego depletion groups are about the same levels in each year.

Figure 11: Annual Profits over individual Years and by ego depletion treatment groups: Averages and 95% confidence intervals



We quantify our results by conducting a series of dummy variable linear regressions using random effects estimators and cluster standard errors at the level of the individuals. The regression results are shown in Table Table 20. In model (1) we regress annual profits on Year where the profit in Year 1 is included in the constant term. In model (2) we introduce dummy variables for the ego depletion treatment variables. The constant reflects the average annual profits of the Low-depletion group in Year 1; the Med-dep and High-dep dummy variables reflect the average annual profit of the two groups in Year 1. In model (3) we include the interaction terms of the Year variable with the ego depletion group dummy variables for the additional learning rate of the High-depletion group and the Med-depletion group.

From the regression results we find that although significant performance improvements were made on average each year, the annual profits made by participants in different ego depletion treatment groups do not significantly different from each other, thus do not support Hypothesis 7. We conclude that learning can be observed from Year 1 through Year 5, however, the learning rate does not significantly varies across treatment groups. Further, the Chow test for the significance of the difference between model (2) and (3) shows no treatment effect of implementing the ego depletion task, (χ^2 (4, N=565) = 1.19, $p = 0.880$).

Table 20: Ego depletion treatment dummy variable regressions for annual profit. (n=565)

	(1) Annual Profit	(2) Annual Profit	(3) Annual Profit
Year	35.12***	35.12***	31.95***
	(3.54)	(3.55)	(5.65)
High-Dep		-3.00	-16.07
		(29.83)	(42.05)
Med-Dep		15.43	9.31
		(25.94)	(35.55)
High-Dep*Year			6.54
			(9.16)
Med-Dep*Year			3.06
			(7.88)
Constant	534.25***	530.04***	536.38***
	(16.28)	(21.43)	(24.89)
chi2	98.19	102.80	105.90
p	0.00	0.00	0.00

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

We then quantify and assess the correlation between CRT groups and ego depletion treatments by conducting a series of linear regressions using random effects estimators and cluster standard errors at the level of the individuals. Results are reported in Table 21. In model (5), we add interaction dummy variables for the CRT groups and ego depletion treatments to examine their joint imposition, using CRT-Level 1 and Low-depletion as base level. In model (6) we include the interaction terms of the Year variable with the CRT groups and the ego depletion group dummy variables. We show the results of the CRT study are robust to this experimental design factor.

Table 21: Dummy variable regressions for annual profit. (n=565)

	(1) Annual Profit	(2) Annual Profit	(3) Annual Profit	(4) Annual Profit	(5) Annual Profit	(6) Annual Profit
CRT-Level1	-123.96***	-194.15***			-154.48***	-226.82***
	(25.53)	(36.09)			(42.13)	(49.90)
CRT-Mixed	-41.81***	-72.39***			-28.86*	-47.71*
	(15.85)	(22.09)			(15.62)	(24.61)
Year		14.28***	35.12***	31.95***	35.12***	14.28***
		(4.23)	(3.55)	(5.65)	(3.56)	(4.26)
CRT-Level1*Year		35.10***				36.17***
		(8.27)				(13.07)
CRT-Mixed*Year		15.29***				9.42
		(5.71)				(6.33)
High-Dep			-3.00	-16.07		
			(29.83)	(42.05)		
Med-Dep			15.43	9.31		
			(25.94)	(35.55)		
High-Dep*Year				6.54		
				(9.16)		
Med-Dep*Year				3.06		
				(7.88)		
CRT-Level1*Med-Dep					67.59	90.98
					(50.59)	(63.61)
CRT-Level1*High-Dep					5.33	-26.46
					(67.18)	(97.47)
CRT-Mixed*Med-Dep					-4.75	-25.45
					(19.53)	(33.12)
CRT-Mixed*High-Dep					-33.66	-48.65
					(30.08)	(38.31)
CRT-Level1*Med-Dep*Year						-11.69
						(15.19)
CRT-Level1*High-Dep*Year						15.89
						(21.14)
CRT-Mixed*Med-Dep*Year						10.35
						(8.37)
CRT-Mixed*High-Dep*Year						7.49
						(9.11)
Constant	672.37***	643.81***	530.04***	536.38***	602.13***	643.81***
	(10.84)	(15.46)	(21.43)	(24.89)	(13.28)	(15.57)
chi2	25.23	125.62	102.80	105.90	117.28	151.06
p	0.00	0.00	0.00	0.00	0.00	0.00

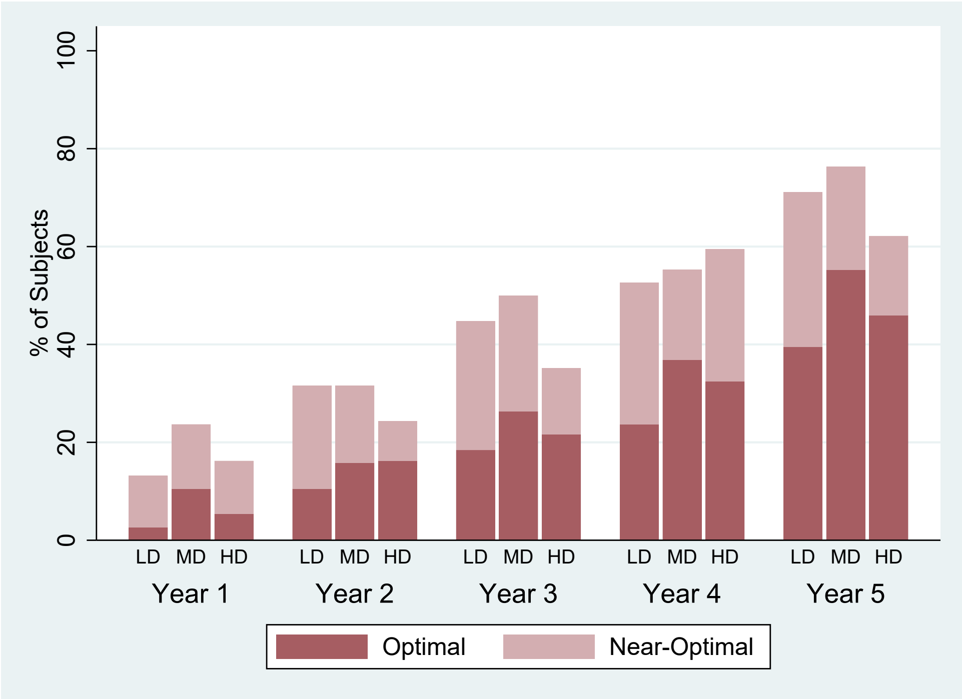
Standard errors in parentheses; Year 1 as baseline.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.4.3 Inventory management policy choices

Figure 12 shows the percentage of participants following optimal and near optimal strategies in each ego depletion groups. In general, we can observe trend in learning to adopt optimal and near optimal policy from Year 1 through 5, however, no significant differences across groups can be found.

Figure 12: Stacked graph of the percentage of participants following optimal and near optimal EOQ constant policies: by Year and ego depletion treatment groups



Result 11. *Participants who have engaged in lower level of ego depletion task did not have higher percentage use of optimal and near-optimal policies in all five years.*

Similarly, we estimate a set of Logit regressions to investigate whether different level of ego depletion tasks is influencing the probability of subjects to adopt EOQ optimal or near optimal policy. The Logit regression results presented in Table 22 show that the level of ego depletion does not have significant impact on policy choices. Finally, we estimate a second sets of Logit regression on the probability of a subject makes Non-EOQ decision. The probabilities of making Non-EOQ decisions are formulated as simple Logit functions of a function of time, accumulated Non-EOQ ordering habit, and different ego depletion groups. The Logit regression results are presented in Table 23, participants engaged in different levels of ego depletion task do not exhibit significantly different Non-EOQ behaviours under both conditions.

Table 22: Logit regression on the probability of choosing optimal or near optimal policy. (n=565)

	(1) Optimal	(2) Optimal	(3) Near-optimal	(4) Near-optimal
$Year$	0.597***	0.604***	0.584***	0.587***
	(0.062)	(0.063)	(0.061)	(0.062)
High-Dep		0.355		-0.152
		(0.381)		(0.331)
Med-Dep		0.624		0.223
		(0.380)		(0.345)
Constant	-3.117***	-3.481***	-2.072***	-2.109***
	(0.269)	(0.356)	(0.236)	(0.298)
χ^2	92.72***	93.81***	91.30***	91.42***
Pr	0.241	0.241	0.432	0.432

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 23: Logit regression on the probability of deviating from an EOQ action

	Panel A: $I_{t-1} < 20$			Panel B: $I_{t-1} \geq 20$		
$NonEOQ_{i,r}$	(1)	(2)	(3)	(4)	(5)	(6)
$Year_r$	-0.176**	-0.182**	-0.306***	-0.484***	-0.487***	-0.998***
	(0.081)	(0.080)	(0.103)	(0.061)	(0.061)	(0.096)
$Month_r$	0.152***	0.154***	0.145***	-0.0844***	-0.0854***	-0.145***
	(0.038)	(0.038)	(0.041)	(0.020)	(0.020)	(0.024)
Medium Depletion		-0.837**	-0.779**		-0.273	-0.0989
		(0.348)	(0.347)		(0.191)	(0.196)
High Depletion		0.137	-0.0525		0.295*	-0.0611
		(0.274)	(0.296)		(0.171)	(0.193)
$NonEOQACC_{i,r-1}$			0.118***			0.305***
			(0.029)			(0.021)
Constant	-3.818***	-3.657***	-3.473***	-1.291***	-1.304***	-0.421
	(0.400)	(0.436)	(0.513)	(0.207)	(0.243)	(0.279)
N	2008	2008	1895	4772	4772	4772
χ^2	26.65***	45.18***	63.62***	75.67***	95.07***	228.5***
$Pr(NonEOQ\_i, r) = 1$	0.0344	0.0344	0.0348	0.0417	0.0417	0.0417

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.5 Conclusion

In Chapter 5 we discussed the experimental results of the ego depletion treatment in the EOQ inventory management. We manipulated different levels of participants' self-regulatory recourse using a variation of the widely accepted letter 'e' task. The treatment aims to study the impact of self-control in a generalization of the EOQ model that incorporates stochastic demand. Recent papers of the (S, s) inventory model (Khaw et al., 2017; Magnani et al., 2016) have found subjects learn to adopt the optimal policy, however, their choices do not persist that they often subsequently deviate from the optimal policy. Both papers speculate the key drivers of such sub-optimal behaviours are related to self-control issues correspond to individuals' bounded rationality. The source of the phenomenon is the substantial cognitive costs arise from responding to the dynamic environment and the noisy feedback.

Ego depletion is a widely used instrument for imposing exogenous load on individuals' limited self-control resources. Preceding studies suggested ego-depletion has negative impacts on peoples' performance on a variety of tasks, while it is not the case in our experiment. The design feature of our EOQ environment follows deterministic demand which gives consistent feedback. Under such condition, people do not make mistakes, our results find no significant ego depletion treatment effects on participants' inventory management performance.

6 Conclusion

The thesis follows a long-term research agenda aiming to examine psychological factors and their impact on EOQ inventory management. The set of experiments in this thesis were designed to serve as the baseline to evaluate the hypotheses that have developed around the more prominent (S, s) inventory model - a foundation model in macroeconomics. Apart from the well-known newsvendor's model, bullwhip effect, etc., the EOQ model is one of the most commonly used model in supply chain management. Despite its prevalence in practice, the EOQ model has drawn less attention in the field of behavioural operations management due to the non-perishable nature of durable goods. However, a good understanding of the EOQ model will improve the efficiency of minimising the total costs incurred in the process of managing durable good. We argue that costs minimisation is essential to prevent the business from significant losses and failures. The current research aims to make contribution to the existing literature that studied EOQ inventory management from a behavioural perspective. Chapter 2 concentrated on model development. Chapter 3, 4 and 5 discussed three experimental research on how cognitive stress, cognitive reflection and ego depletion affect inventory management decisions using the EOQ inventory management model.

In Chapter 3, the first aspect of our experimental design is the presence of an additional task that competes for the participants' short-term memory resources, we call this our "High" treatment. Our "Low" treatment doesn't involve this competing task. The second aspect of our experimental design involves an intervention, the "Zero-Only" treatment limits the complexity of the inventory policy choice set. It removes the possibility of violating the optimal inventory policy by forbidding participants from ordering when there is a positive level of inventory. On the contrary, participants can order additional inventory each month regardless of the current inventory level in the "Unrestricted" treatment. Our results suggest that increases in cognitive load negatively impact participants' performance. However, these negative impacts occur predominantly when participants first face the inventory decision problem. While average performance is not statistically different, we note that only in the Zero Only-Low treatment cell do we observe the majority of participants eventually learn to use the optimal EOQ policy.

In Chapter 4, our finite horizon EOQ setting does not involve uncertainty nor strategic considerations. These premises ensure the evaluations of the relationship between the cognitive state of reflection, gender, and performance in EOQ inventory management without any influences of individual risk attitudes, nor the consideration of strategic uncertainty and coordination. Firstly, we observe that participants with greater performance in the CRT task earn more and make more optimal decisions. Male participants also outperform females in both tasks. Also, we note that these differences do not persist, the performance gap closes along repetitions of our experimental rounds. Further, to address the gap of gender performance, we proved in a multivariate analysis that control for both gender and CRT performance that the such gap in decision quality is not gender drive; differences in CRT performance are the main driver of earnings differences.

In Chapter 5, we discussed the results of the ego depletion treatment we have implemented alongside the CRT treatment in the same experiment. The treatment variable is the level of temporary depletion of self-regulatory capacity by an initial act of self-control. We have improved the design of previous self-regulation research by incentivise the ego depletion task. Participants tend not to deliberately make random selections which require no self-control. Also, we switched the traditional pen and paper approach to a computerized task, so that the process of monitoring and payment calculation becomes instant. The design of our EOQ environment follows deterministic demand which gives consistent feedback. Under such static condition, our results show there is no obvious treatment effect on participants' self-regulation ability.

Some natural next steps are to explore how the choice set complexity, self-control depletion and corresponding framing impacts decision-making in the other previously raised inventory management frameworks such as the newsvendor problem, (S, s) inventory management, and multi-echelon supply chains. We believe our findings on performance improvements from external decision support offers much scope for future research.

References

Abeler, J., & Jäger, S. (2015). Complex tax incentives. *American Economic Journal: Economic Policy*, *7*(3), 1–28.

Allred, S., Duffy, S., & Smith, J. (2016). Cognitive load and strategic sophistication. *Journal of Economic Behavior & Organization*, *125*, 162–178.

Allred, S. R., Crawford, L. E., Duffy, S., & Smith, J. (2016). Working memory and spatial judgments: Cognitive load increases the central tendency bias. *Psychonomic bulletin & review*, *23*(6), 1825–1831.

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*(5), 1252.

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2013). Who is 'behavioral'? cognitive ability and anomalous preferences. *Journal of the European Economic Association*, *11*(6), 1231–1255.

Bergman, O., Ellingsen, T., Johannesson, M., & Svensson, C. (2010). Anchoring and cognitive ability. *Economics Letters*, *107*(1), 66–68.

Bode, C., & Wagner, S. M. (2015). Structural drivers of upstream supply chain complexity and the frequency of supply chain disruptions. *Journal of Operations Management*, *36*, 215–228.

Bolton, G. E., & Katok, E. (2008). Learning by doing in the newsvendor problem: A laboratory investigation of the role of experience and feedback. *Manufacturing & Service Operations Management*, *10*(3), 519–538.

Bostian, A. A., Holt, C. A., & Smith, A. M. (2008). Newsvendor "pull-to-center" effect: Adaptive learning in a laboratory experiment. *Manufacturing & Service Operations Management*, *10*(4), 590–608.

Brañas-Garza, P., Garcia-Muñoz, T., & González, R. H. (2012). Cognitive effort in the beauty contest game. *Journal of Economic Behavior & Organization*, *83*(2), 254–260.

Brañas-Garza, P., Kujal, P., & Lenkei, B. (2015). *Cognitive Reflection Test: Whom, how, when* (MPRA Paper No. 68049).

Buser, T., & Peter, N. (2012). Multitasking. *Experimental Economics*, *15*(4), 641–655.

Campitelli, G., & Gerrans, P. (2014). Does the cognitive reflection test measure cognitive reflection? a mathematical modeling approach. *Memory & cognition*, *42*(3), 434–447.

Caplin, A., Dean, M., & Martin, D. (2011). Search and satisficing. *American Economic Review*, *101*(7), 2899–2922.

Carpenter, J., Graham, M., & Wolf, J. (2013). Cognitive ability and strategic sophistication. *Games and Economic Behavior*, *80*, 115–130.

Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree - an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, *9*, 88–97.

Chen, K.-Y., & Wu, D. (2017). Learning under the deterministic economic order quantity problem. *University of Texas - Arlington*.

Croson, R., Donohue, K., Katok, E., & Sterman, J. (2014). Order stability in supply chains: Coordination risk and the role of coordination stock. *Production and Operations Management*, *23*(2), 176–196.

Cueva, C., Iturbe-Ormaetxe, I., Mata-Pérez, E., Ponti, G., Sartarelli, M., Yu, H., & Zhukova, V. (2016). Cognitive (ir) reflection: New experimental evidence. *Journal of Behavioral and Experimental Economics*, *64*, 81–93.

de Vericourt, F., Jain, K., Bearden, J. N., & Filipowicz, A. (2013). Sex, risk and the newsvendor. *Journal of Operations Management*, *31*(1-2), 86–92.

Deck, C., & Jahedi, S. (2015). The effect of cognitive load on economic decision making: A survey and new experiments. *European Economic Review*, *78*, 97–119.

DeWall, C. N., Baumeister, R. F., Stillman, T. F., & Gailliot, M. T. (2007). Violence restrained: Effects of self-regulation and its depletion on aggression. *Journal of Experimental social*

psychology, *43*(1), 62–76.
Duffy, S., & Smith, J. (2014). Cognitive load in the multi-player prisoner's dilemma game: Are there brains in games? *Journal of Behavioral and Experimental Economics*, *51*, 47–56.
Feng, T., Keller, L. R., & Zheng, X. (2011). Decision making in the newsvendor problem: A cross-national laboratory study. *Omega*, *39*(1), 41–50.
Fennis, B. M., Janssen, L., & Vohs, K. D. (2008). Acts of benevolence: A limited-resource account of compliance with charitable requests. *Journal of Consumer Research*, *35*(6), 906–924.
Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, *19*(4), 25–42.
Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. (2010). Ego depletion and the strength model of self-control: a meta-analysis. *Psychological bulletin*, *136*(4), 495.
Harper, A. (2008). Psychometric tests are now a multi-million-pound business: What lies behind a coach's decision to use them. *International Journal of Evidence Based Coaching and Mentoring*, *2*, 40–51.
Harris, F. W. (1913). How Many Parts to Make At Once. *Factory, the magazine of management*, *10*(2), 135–136. Retrieved from http://www.jstor.org/stable/170962 doi: 10.2307/170962
Hinson, J. M., Jameson, T. L., & Whitney, P. (2003). Impulsive decision making and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(2), 298.
Hoppe, E. I., & Kusterer, D. J. (2011). Behavioral biases and cognitive reflection. *Economics Letters*, *110*(2), 97–100.
Khaw, M. W., Stevens, L., & Woodford, M. (2017). Discrete adjustment to a changing environment: Experimental evidence. *Journal of Monetary Economics*, *91*, 88–103.
Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of behavioral decision making*, *25*(4), 361–381.
Lleras, J. S., Masatlioglu, Y., Nakajima, D., & Ozbay, E. Y. (2017). When more is less: Limited consideration. *Journal of Economic Theory*, *170*, 70–85.
Long, X., & Nasiry, J. (2014). Prospect theory explains newsvendor behavior: The role of reference points. *Management Science*, *61*(12), 3009–3012.
Magnani, J., Gorry, A., & Oprea, R. (2016). Time and state dependence in an (S, s) decision experiment. *American Economic Journal: Macroeconomics*, *8*(1), 285–310.
Mangan, J., & Christopher, M. (2005). Management development and the supply chain manager of the future. *The International Journal of Logistics Management*, *16*(2), 178–191.
Masatlioglu, Y., Nakajima, D., & Ozbay, E. Y. (2012). Revealed attention. *American Economic Review*, *102*(5), 2183–2205.
Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81.
Moritz, B., Siemsen, E., & Kremer, M. (2014). Judgmental forecasting: Cognitive reflection and decision speed. *Production and Operations Management*, *23*(7), 1146–1160.
Moritz, B. B., Hill, A. V., & Donohue, K. L. (2013). Individual differences in the newsvendor problem: Behavior and cognitive reflection. *Journal of Operations Management*, *31*(1-2), 72–85.
Morse, G. (2016). Designing a bias-free organization. *Harvard business review*, *94*(7), 15.
Muraven, M. (2008). Prejudice as self-control failure 1. *Journal of Applied Social Psychology*, *38*(2), 314–333.
Narayanan, A., & Moritz, B. B. (2015). Decision making and cognition in multi-echelon supply chains: An experimental study. *Production and Operations Management*, *24*(8), 1216–1234.

Newman, D. A., & Lyon, J. S. (2009). Recruitment efforts to reduce adverse impact: Targeted recruiting for personality, cognitive ability, and diversity. *Journal of Applied Psychology*, *94*(2), 298.

Neys, W. D. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological science*, *17*(5), 428–433.

Niranjan, T. T., Wagner, S. M., & Bode, C. (2011). An alternative theoretical explanation and empirical insights into overordering behavior in supply chains. *Decision Sciences*, *42*(4), 859–888.

Oechssler, J., Roider, A., & Schmitz, P. W. (2009). Cognitive abilities and behavioral biases. *Journal of Economic Behavior & Organization*, *72*(1), 147–152.

Oracle. (2018). *What is erp?* Retrieved 2018-09-12, from https://www.oracle.com/applications/erp/what-is-erp.html

Pan, J., Shachat, J., & Wei, S. (2019). *Cognitive stress and learning Economic Order Quantity (EOQ) inventory management: An experimental investigation* (MPRA Paper No. 86221).

Pease, A., & Pease, B. (2001). *Why Men Don't Listen and Women Can't Read Maps.* London: Orion.

Pease, A., & Pease, B. (2003). *Why Men Can Only Do One Thing at a Time and Women Never Stop Talking.* London: Orion.

Pennycook, G., Cheyne, J. A., Koehler, D. J., & Fugelsang, J. A. (2016). Is the cognitive reflection test a measure of both reflection and intuition? *Behavior Research Methods*, *48*(1), 341–348.

Primi, C., Morsanyi, K., Chiesi, F., Donati, M. A., & Hamilton, J. (2016). The development and testing of a new version of the cognitive reflection test applying item response theory (irt). *Journal of Behavioral Decision Making*, *29*(5), 453–469.

Roch, S. G., Lane, J. A., Samuelson, C. D., Allison, S. T., & Dent, J. L. (2000). Cognitive load and the equality heuristic: A two-stage model of resource overconsumption in small groups. *Organizational Behavior and Human Decision Processes*, *83*(2), 185–212.

Ruderman, M. N., Clerkin, C., & Deal, J. J. (2017). The long-hours culture. In C. L. Cooper & M. P. Leiter (Eds.), *The routledge companion to wellbeing at work* (pp. 207 –221). Taylor & Francis.

Rudi, N., & Drake, D. (2014). Observation bias: The impact of demand censoring on newsvendor level and adjustment behavior. *Management Science*, *60*(5), 1334–1345.

Rydval, O. (2012). The causal effect of cognitive abilities on economic behavior: evidence from a forecasting task with varying cognitive load. *CERGE-EI Working Paper Series*(457).

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological bulletin*, *124*(2), 262.

Schwarz, L. B. (1972). Economic order quantities for products with finite demand horizons. *A I I E Transactions*, *4*, 234–237.

Schweitzer, M. E., & Cachon, G. P. (2000). Decision bias in the newsvendor problem with a known demand distribution: Experimental evidence. *Management Science*, *46*(3), 404–420.

Shiv, B., & Fedorikhin, A. (1999). Heart and mind in conflict: The interplay of affect and cognition in consumer decision making. *Journal of Consumer Research*, *26*(3), 278–292.

Sprenger, A. M., Dougherty, M., Atkins, S. M., Franco-Watkins, A. M., Thomas, R., Lange, N., & Abbs, B. (2011). Implications of cognitive load for hypothesis generation and probability judgment. *Frontiers in psychology*, *2*, 129.

Stangl, T., & Thonemann, U. W. (2017). Equivalent inventory metrics: A behavioral perspective. *Manufacturing & Service Operations Management*, *19*(3), 472–488.

Stanovich, K. E., & West, R. F. (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and brain sciences*, *23*(5), 645–665.

Tokar, T., Aloysius, J. A., & Waller, M. A. (2012). Supply chain inventory replenishment: The debiasing effect of declarative knowledge. *Decision Sciences*, *43*(3), 525–546.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & cognition*, *39*(7), 1275.

Tyler, J. M. (2008). In the eyes of others: Monitoring for relational value cues. *Human Communication Research*, *34*(4), 521–549.

Wang, C. X., & Webster, S. (2009). The loss-averse newsvendor problem. *Omega*, *37*(1), 93–105.

Wheeler, S. C., Briñol, P., & Hermann, A. D. (2007). Resistance to persuasion as self-regulation: Ego-depletion and its effects on attitude change processes. *Journal of Experimental Social Psychology*, *43*(1), 150–156.

A Appendix for Chapter 3

A.1 Experiment Instructions and Interface

Instructions for different treatments are presented as the texts/sentences in italics and square brackets below.

A.1.1 Instruction Page

Welcome

Welcome to today's experiment. Please read the following instructions carefully as they are directly relevant to how much money you will earn today. Please do not communicate with other people during the experiment. Please note that **you are not permitted to use pen and paper or a mobile phone**. Please kindly switch your mobile phone off or put it on silent mode. Students causing a disturbance will be asked to leave the room. You will enter all of your decisions in today's experiment using only the computer mouse. Please do not attempt to use the keyboard or remove the keyboard cover. The information displayed on your computer monitor is private and specific to you. All monetary amounts in today's experiment are expressed as experimental currency units (ECU). The conversion rate for ECU and GBP is 300 ECU = £1 cash payment. Your payment will be rounded up to the nearest ten pence.

If you have any questions at any point during today's session, please raise your hand and one of the monitors will come to help.

Task

In today's experiment, you will be making **inventory management decisions** for an enterprise called 'S-Store'. S-Store sells coffee makers. You will perform this role for a sequence of 6 years. Every month you will decide how many coffee makers to order from the coffee maker supplier. Your earnings in this experiment will be proportional to the total profitability of S-Store. S-store will sell a new coffee maker model every year. Thus in the first month of a year your inventory always starts from zero. Further, any coffee makers remaining in inventory at the end of month 12 will be disposed of. To summarise, you will be making 12 monthly decisions for a year, and you will do this for 6 years in total.

You will have up to **4 minutes** to complete your task for each year. Year 1 is a practice round, and you will have up to 7 minutes to complete the task for this year. You should use this as an opportunity to familiarize yourself with the software and decision tasks. If you don't finish within the time allowed, the computer will automatically execute the remaining month(s) sales with the existing inventory. You will not be able to add inventory. A 'wait page' displays automatically if you spend less than the allowed time in a year. You will only be able to proceed to the next year when the remaining time runs out.

Before the decision making portion of the experiment begins, there will be a **Quiz** consisting of 7 simple questions to check your understanding of the task. Please answer the questions carefully. If you missed 3 or more questions, you would be asked review the correct answers before you can proceed to the task.

[The following italic texts are additional for treatments with High Cognitive Loads]

PIN

In addition to the task, you will be given a 7-digit PIN at the beginning of each year. The PIN is case sensitive, and consisting of numbers, uppercase and lowercase letters. You will have 15 seconds to remember the PIN. This is your KEY to unlock an account which contains an extra reward of 300 ECU. You can open the account at the end of each year by correctly entering the PIN. You will only have one attempt to correctly enter the pin to claim this extra reward.

Payment

Year 0 is a practice round, and you will receive no earnings from your decisions in this year. For **Years 1 through 5**, your earnings will accumulate across years. At the end of the experiment you will be paid £5 show-up fee and your accumulated earnings, converted to Pounds. Note, negative profit may occur if poor coffee maker ordering decisions are made. To ensure that no one will leave the experiment with a payment less than £5, a negative total profit made in Year 1 to Year 5 will be treated as 0 earnings.

A.1.2 Background Information

[The following Background Information section shows up on every decision page.]

Your Role:

S-Store is open 360 days per year. You are the inventory manager for S-Store. In your role, you will control S-Store's inventory level which determines the store's total profits.

We now explain how S-Stores, and correspondingly you, earns profit. While we are explaining how the calculations are made, during the decision tasks the computer will carry out these calculations and report the results to you.

S-Store sells coffee makers at a price of **7 ECU** per unit. S-Store can sell up to **10 coffee makers per month**. A coffee maker can only be sold if there is a unit held in inventory. If you hold 10 or more units in inventory at the start of the month, S-Store will sell 10 coffee makers that month. However, if there are less than 10 units held in inventory at the start of the month then S-Store will only sell that amount. For example, if there are 2 units held in inventory at the beginning of a month then S-Store only sells 2 units that month. *[(For Zero Only treatment only) You can only place an order when the current month's opening inventory is 0. For example, if the current month's opening inventory is 3 units, you cannot place an order this month, S-Store only sells 3 units this month.]* S-Store's sales revenue for a month is calculated as follows:

$$\text{Sales revenue} = 7 \text{ ECU} * \text{Number of units sold.}$$

Your job is to manage the store's inventory levels by each month choosing an inventory order. Prior to the start of each month you can order coffee makers from the supplier to add to the inventory. Your inventory management determines the S-Store's total costs. S-Store pays two types of costs. One is the **ordering cost**. Every time you order a positive amount you have to pay an order cost. This ordering cost is **45 ECU**, and does not depend upon the size of the order. If you order zero coffee makers then you do not pay the 45 ECU ordering cost. Holding coffee makers in inventory is costly so S-Store pays a monthly **inventory holding cost**. S-Store pay's monthly inventory holding cost is based on the average number of coffee makers held in inventory multiplied by the per unit monthly inventory holding cost of **1 ECU**. This is calculated as follows:

$$\text{Inventory holding costs} = 1 \text{ ECU} * (\text{Opening inventory} + \text{Order Quantity} + \text{Closing inventory})/2.$$

Calculation of S-Store's profits

$$\text{Profits} = \text{Sales revenue - Ordering costs - Inventory holding costs}$$

Your monthly earnings are equal to S-Store's monthly profits.

Examples:

1. Alice's closing inventory of last month is 20 units, she placed an order of 0 units in this month.

The demand for each month is 10 units.
She made sales of 10 units.
Her closing inventory of this month is $20 - 10 = 10$ units.
Her profit in this month is equal to: $7 * 10 - 0 - 1 * (20 + 0 + 10)/2 = 55$.

2. Alice's closing inventory of last month is 4 units, she placed an order of 5 units in this month.
   The demand for each month is 10 units.
   She only made sales of 9 units. Her closing inventory of this month is 0 units. Her profit in this month is equal to: $7 * 9 - 45 - 1 * (4 + 5 + 0)/2 = 13.5$.

A.1.3 Multiple Choice Questions prior to Decision Task

There are a couple of questions for you before the task, please use the information:

The demand for each month is 10 units.
Price of each coffee maker is 7.
Ordering cost is 45 per order.
Monthly inventory holding cost is 1 per unit.

Question 1 of 7
If the inventory level was 5 and you ordered 0 units. How many units will you SELL this month?

A 0
B 5
C 10
D 15

Question 2 of 7
If the inventory level was 0 and you ordered 15 units. How many units will you SELL this month?

A 0
B 5
C 10
D 15

Question 3 of 7
If you made sales of 10 units. What will be your SALES REVENUE this month?

A 0
B 10
C 25
D 70

Question 4 of 7
If you ordered 0 units. What will be your ORDERING COST this month?

A 0
B 1
C 45
D 70

Question 5 of 7
If you ordered 1 unit. What will be your ORDERING COST this month?

A 0
B 1
C 45
D 70

Question 6 of 7

If the inventory level was 0 and you ordered 10 units. You made sales of 10 units. What will be your HOLDING COST this month?

A 0
B 1
C 5
D 10

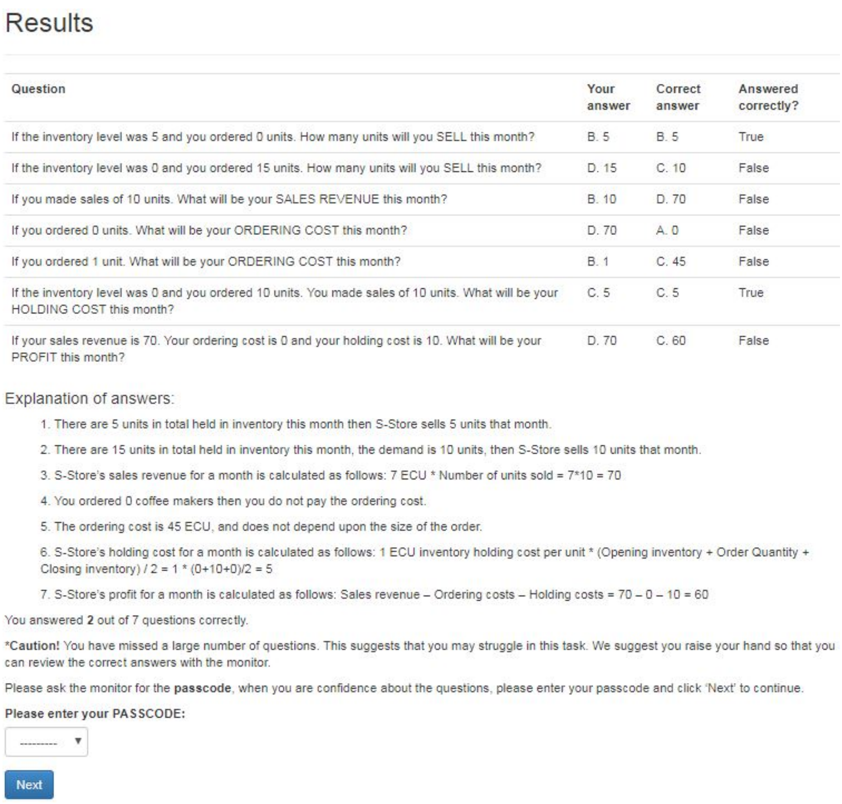
Question 7 of 7

If your sales revenue is 70. Your ordering cost is 0 and your holding cost is 10. What will be your PROFIT this month?

A 15
B 25
C 60
D 70

Figure 13 shows the result page of the multiple choice questions when participants had given more than 2 incorrect answers. Under such circumstances, they had to raise their hands to go through incorrectly answered questions with a monitor in order to obtain a passcode to proceed to the decision tasks.

Figure 13: Result Page of the Multiple Choice Questions



Results

Question	Your answer	Correct answer	Answered correctly?
If the inventory level was 5 and you ordered 0 units. How many units will you SELL this month?	B. 5	B. 5	True
If the inventory level was 0 and you ordered 15 units. How many units will you SELL this month?	D. 15	C. 10	False
If you made sales of 10 units. What will be your SALES REVENUE this month?	B. 10	D. 70	False
If you ordered 0 units. What will be your ORDERING COST this month?	D. 70	A. 0	False
If you ordered 1 unit. What will be your ORDERING COST this month?	B. 1	C. 45	False
If the inventory level was 0 and you ordered 10 units. You made sales of 10 units. What will be your HOLDING COST this month?	C. 5	C. 5	True
If your sales revenue is 70. Your ordering cost is 0 and your holding cost is 10. What will be your PROFIT this month?	D. 70	C. 60	False

Explanation of answers:

1. There are 5 units in total held in inventory this month then S-Store sells 5 units that month.
2. There are 15 units in total held in inventory this month, the demand is 10 units, then S-Store sells 10 units that month.
3. S-Store's sales revenue for a month is calculated as follows: 7 ECU * Number of units sold = 7*10 = 70
4. You ordered 0 coffee makers then you do not pay the ordering cost.
5. The ordering cost is 45 ECU, and does not depend upon the size of the order.
6. S-Store's holding cost for a month is calculated as follows: 1 ECU inventory holding cost per unit * (Opening inventory + Order Quantity + Closing inventory) / 2 = 1 * (0+10+0)/2 = 5
7. S-Store's profit for a month is calculated as follows: Sales revenue – Ordering costs – Holding costs = 70 – 0 – 10 = 60

You answered **2** out of 7 questions correctly.

Caution! You have missed a large number of questions. This suggests that you may struggle in this task. We suggest you raise your hand so that you can review the correct answers with the monitor.

Please ask the monitor for the **passcode**, when you are confidence about the questions, please enter your passcode and click 'Next' to continue.

Please enter your PASSCODE:

Next

A.1.4 Decision Tasks

Prior to each year's decision tasks, a mini-instruction page appears. Figure 14 is an example with PIN task. For treatments with high cognitive loads, the pin page follows (Figure 15).

An example of the ordering decision page is shown in Figure 16. Participants move the slider to enter their decision of order quantity for each month. Order quantities, costs, and profits of

Figure 14: An example of Instruction Page with PIN task

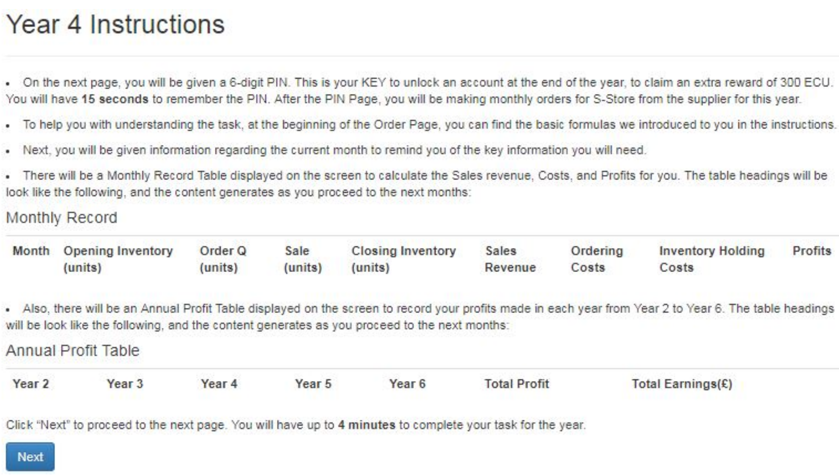
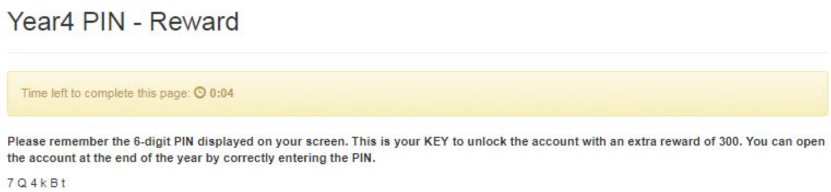


Figure 15: PIN Page prior to Ordering Page



previous months are also displayed on the page. If participants completed the year's decision task within 4 minutes, they had to wait until the end of 4 minutes.

They were then prompted to enter the PIN (Figure 17), followed by the end of the year result page (Figure 18).

A.2 Post-Experimental Survey, Demographics and Summary Statistics of Participants

Participants were asked to fill a simple questionnaire at the end of the experiment for us to collect some demographic information.

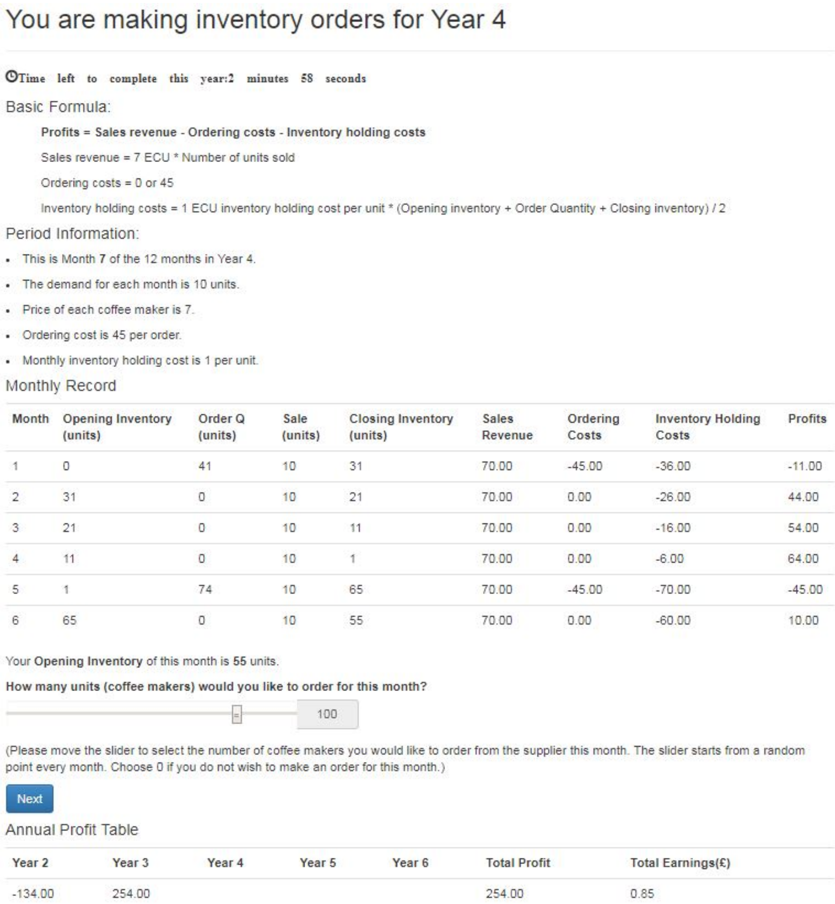
The following are some summary statistics of the participants.

Table 24: Demographics in Participants

Age (mean)	25.6
Gender (% female)	65%
Education (%Undergraduate)	51%

One can observe that 37% of the participants are from Social Science & Management, among which they may have training in operations management or have been exposed to the EOQ model before.

Figure 16: Ordering Page



You are making inventory orders for Year 4

Time left to complete this year:2 minutes 58 seconds

Basic Formula:

Profits = Sales revenue - Ordering costs - Inventory holding costs

Sales revenue = 7 ECU * Number of units sold

Ordering costs = 0 or 45

Inventory holding costs = 1 ECU inventory holding cost per unit * (Opening inventory + Order Quantity + Closing inventory) / 2

Period Information:

- This is Month **7** of the 12 months in Year 4.
- The demand for each month is 10 units.
- Price of each coffee maker is 7.
- Ordering cost is 45 per order.
- Monthly inventory holding cost is 1 per unit.

Monthly Record

Month	Opening Inventory (units)	Order Q (units)	Sale (units)	Closing Inventory (units)	Sales Revenue	Ordering Costs	Inventory Holding Costs	Profits
1	0	41	10	31	70.00	-45.00	-36.00	-11.00
2	31	0	10	21	70.00	0.00	-26.00	44.00
3	21	0	10	11	70.00	0.00	-16.00	54.00
4	11	0	10	1	70.00	0.00	-6.00	64.00
5	1	74	10	65	70.00	-45.00	-70.00	-45.00
6	65	0	10	55	70.00	0.00	-60.00	10.00

Your **Opening Inventory** of this month is **55** units.

How many units (coffee makers) would you like to order for this month?

100

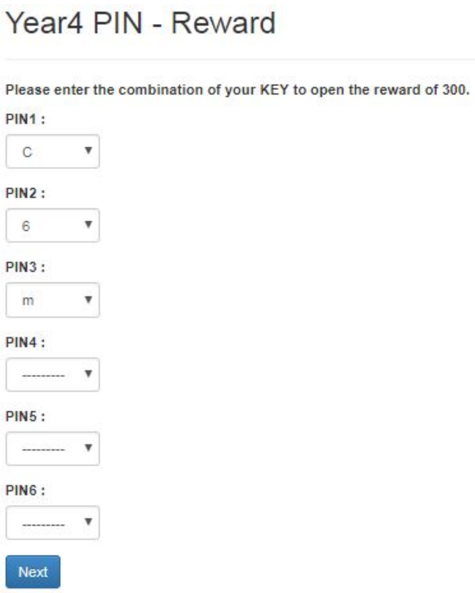
(Please move the slider to select the number of coffee makers you would like to order from the supplier this month. The slider starts from a random point every month. Choose 0 if you do not wish to make an order for this month.)

Next

Annual Profit Table

Year 2	Year 3	Year 4	Year 5	Year 6	Total Profit	Total Earnings(£)
-134.00	254.00				254.00	0.85

Figure 17: Enter the PIN Page



Year4 PIN - Reward

Please enter the combination of your KEY to open the reward of 300.

PIN1 : C

PIN2 : 6

PIN3 : m

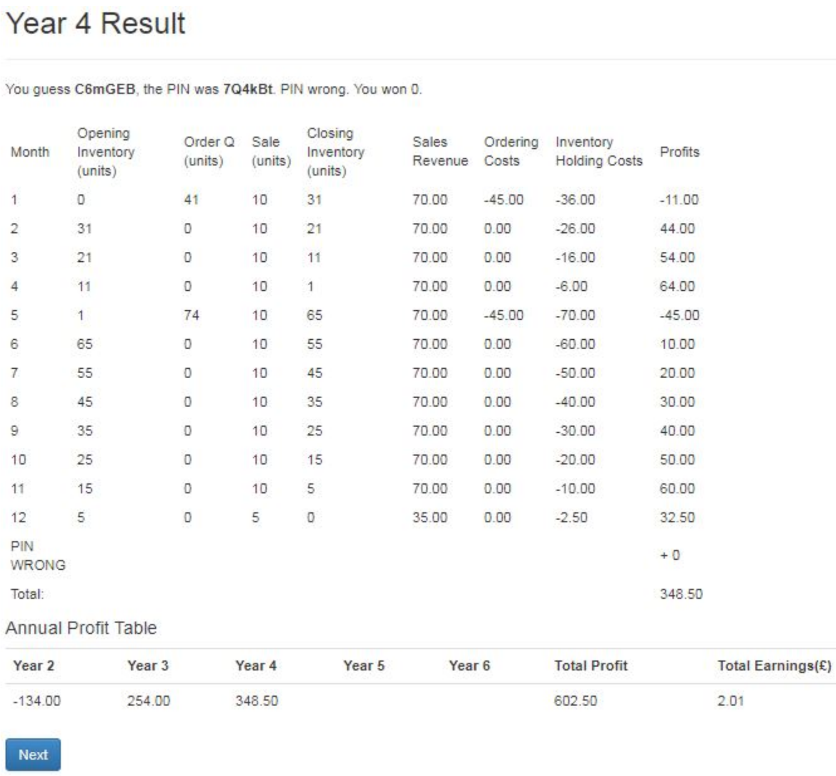
PIN4 : ---------

PIN5 : ---------

PIN6 : ---------

Next

Figure 18: End of the Year Result Page



Year 4 Result

You guess **C6mGEB**, the PIN was **7Q4kBt**. PIN wrong. You won 0.

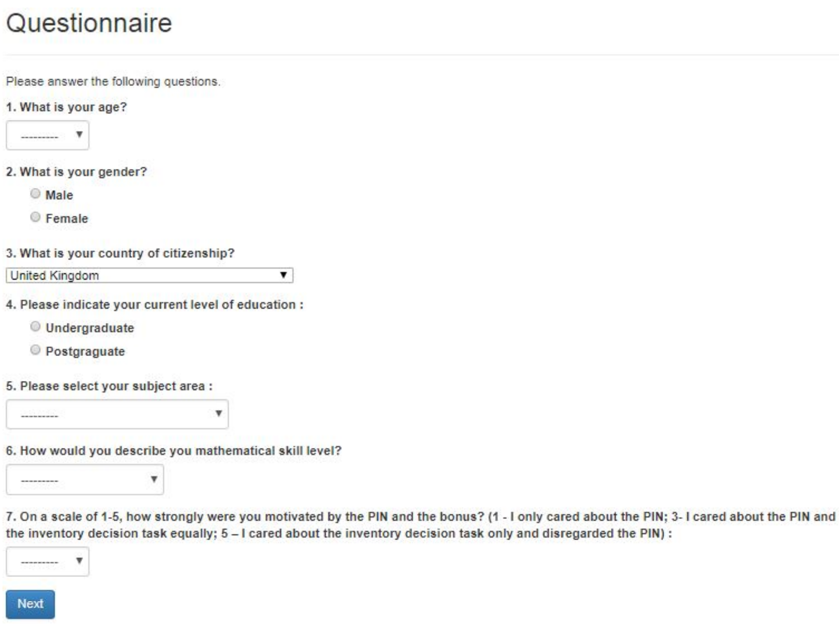
Month	Opening Inventory (units)	Order Q (units)	Sale (units)	Closing Inventory (units)	Sales Revenue	Ordering Costs	Inventory Holding Costs	Profits
1	0	41	10	31	70.00	-45.00	-36.00	-11.00
2	31	0	10	21	70.00	0.00	-26.00	44.00
3	21	0	10	11	70.00	0.00	-16.00	54.00
4	11	0	10	1	70.00	0.00	-6.00	64.00
5	1	74	10	65	70.00	-45.00	-70.00	-45.00
6	65	0	10	55	70.00	0.00	-60.00	10.00
7	55	0	10	45	70.00	0.00	-50.00	20.00
8	45	0	10	35	70.00	0.00	-40.00	30.00
9	35	0	10	25	70.00	0.00	-30.00	40.00
10	25	0	10	15	70.00	0.00	-20.00	50.00
11	15	0	10	5	70.00	0.00	-10.00	60.00
12	5	0	5	0	35.00	0.00	-2.50	32.50
PIN WRONG								+ 0
Total:								348.50

Annual Profit Table

Year 2	Year 3	Year 4	Year 5	Year 6	Total Profit	Total Earnings(£)
-134.00	254.00	348.50			602.50	2.01

Next

Figure 19: Post-Experimental Survey



Questionnaire

Please answer the following questions.

1. What is your age?

----

2. What is your gender?

- Male
- Female

3. What is your country of citizenship?

United Kingdom

4. Please indicate your current level of education :

- Undergraduate
- Postgraguate

5. Please select your subject area :

----

6. How would you describe you mathematical skill level?

----

7. On a scale of 1-5, how strongly were you motivated by the PIN and the bonus? (1 - I only cared about the PIN; 3- I cared about the PIN and the inventory decision task equally; 5 – I cared about the inventory decision task only and disregarded the PIN) :

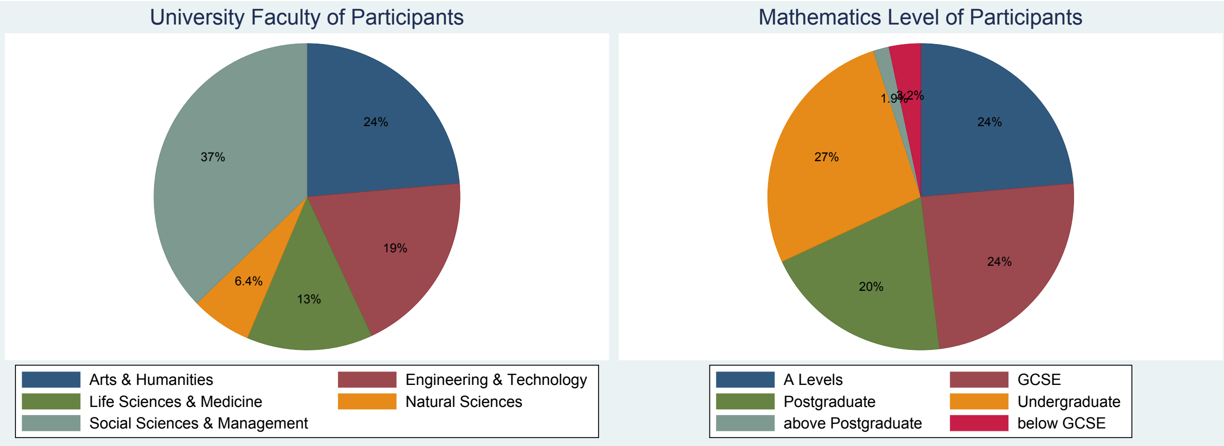
Next

B Appendix for Chapter 4 and Chapter 5

B.1 Experiment Instructions and Interface

Instructions for different treatments are presented as the texts/sentences in italics and square brackets below.

Figure 20: Distribution of University Programs Participants study and Mathematics levels of participants (self reported)



(a) Distribution of University Programs Participants study

(b) Mathematics levels of participants

Table 25: Regression on PIN and demographic information

	(1) Annual Profit	(2) Annual Profit
Year 1	-129.08***	-130.78***
	(19.30)	(18.84)
Year 2	-71.18***	-70.65***
	(16.84)	(16.90)
Year 3	-31.00**	-30.90**
	(13.76)	(13.23)
Year 4	-13.95	-13.74
	(12.28)	(11.89)
Unrestricted	-22.32**	-33.26***
	(11.22)	(12.56)
High·Correct PIN	27.26*	19.09
	(15.96)	(16.13)
Age		-2.42***
		(0.78)
Male		16.14
		(11.88)
Postgrad		-7.44
		(12.40)
STEM		8.66
		(13.73)
Math level		-5.98
		(5.20)
Constant	417.18***	505.75***
	(15.61)	(33.00)
N	385	385
R^2	0.18	0.22

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

B.1.1 Instruction Page

Welcome

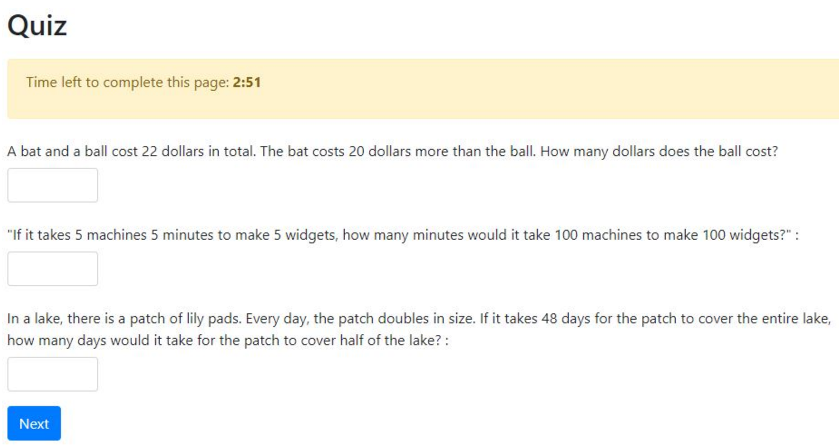
Welcome to today's experiment. Please read the following instructions carefully as they are directly relevant to how much money you will earn today. Please do not communicate with other people during the experiment. Please note that **you are not permitted to use pen and paper or a mobile phone**. Please kindly switch your mobile phone off or put it on silent mode. Students causing a disturbance will be asked to leave the room. The information displayed on your computer monitor is private and specific to you. All monetary amounts in today's experiment are expressed as experimental currency units (ECU). The conversion rate for ECU and GBP is 450 ECU = £1 cash payment. Your payment will be rounded up to the nearest ten pence.

There are **three tasks** in total, a '**Quiz**', a '**Letter Task**' and an '**Inventory Task**'. At the end of the experiment you will be paid £5 show-up fee and your accumulated earnings, converted to Pounds. If you have any questions at any point during today's session, please raise your hand and one of the monitors will come to help.

Quiz

You will have up to 3 minutes to answer 3 short questions. Each correct answer gains 300 ECU. When you are ready please click "Next" to begin. (The interface of Quiz is shown as Figure 21.)

Figure 21: CRT assessment interface



Letter Task

Please **delete** words with the occurrences of the letter **'e'**; otherwise choose keep.

The task involving 150 words *[50 words for Medium and High Depletion treatment]* with 10 words on each page, you will have up to **50 seconds** to complete each page. Your payment on this task will be calculated based on your accuracy of completion. Each correct answer gains 15 ECU. When you are ready please click "Next" to begin.

Examples:

1. 'Apple' would be deleted;
2. 'School' would be kept.

[The followings only occur for Medium and High Depletion treatment. High Depletion treatment is indicated with square bracket]

Rule change

*Please **delete** words with the occurrences of the letter '**e**'. But **keep** the word if the letter '**e**' occurs **next to another vowel** or **one letter away from another vowel** [or if it is also an **adjective**].*

**Note that the letters A, E, I, O, and U are called vowels; [an "adjective" is a word that describes a noun or pronoun, "big", "boring", "purple", and "obvious" are all adjectives.]*

*The task involving 100 words with 10 words on each page, you will have up to **50 seconds** to complete each page. Your payment on this task will be calculated based on your accuracy of completion. Each correct answer gains 15 ECU. When you are ready please click "Next" to begin.*

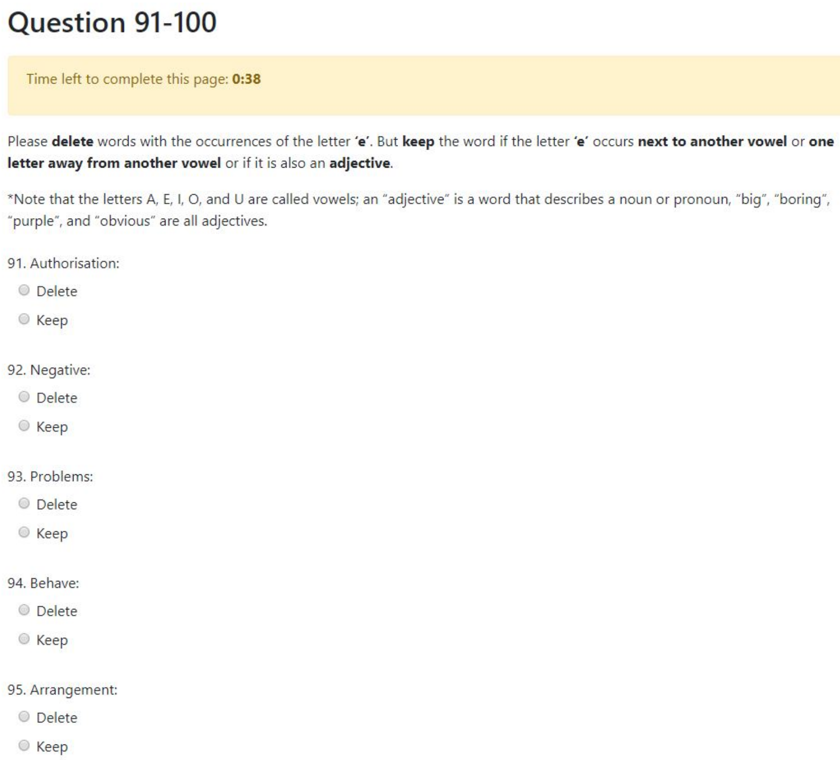
***Examples**:*

1. *'Apple' would be deleted;*
2. *'School' would be kept.*
3. *'Read' would be kept;*
4. *'Towel' would be kept;*

[5. 'Excellent' (adj.) would be kept.]

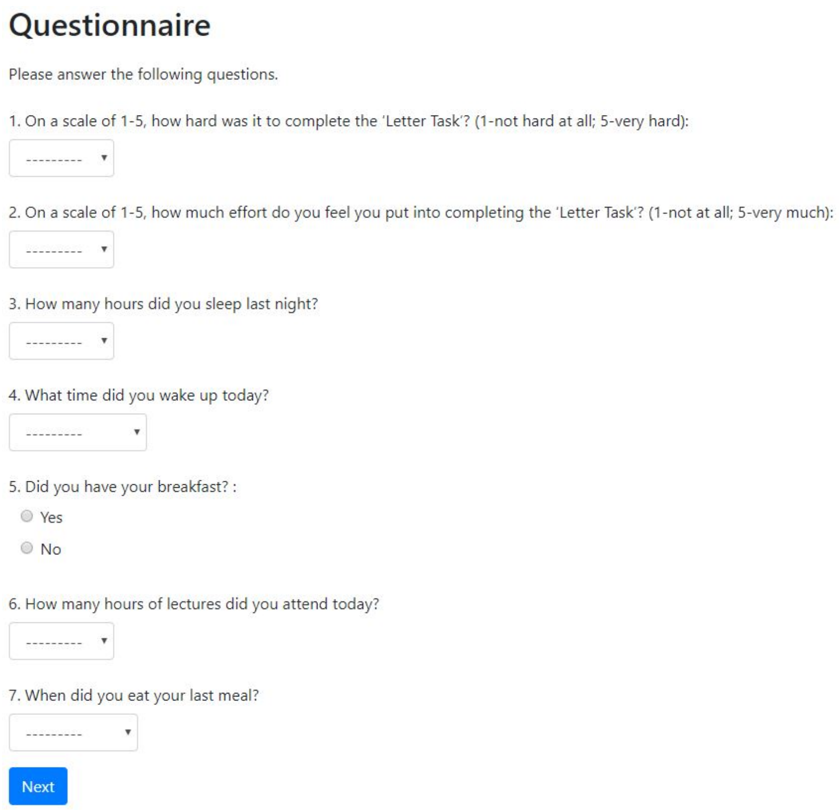
An example of the interface of the Letter Task can be found in Figure 22.

Figure 22: An example of the interface of the Letter Task (High Depletion)



The questionnaire after the Letter Task is shown in Figure 23.

Figure 23: Questionnaire after Letter Task



Questionnaire

Please answer the following questions.

1. On a scale of 1-5, how hard was it to complete the 'Letter Task'? (1-not hard at all; 5-very hard):

----

2. On a scale of 1-5, how much effort do you feel you put into completing the 'Letter Task'? (1-not at all; 5-very much):

----

3. How many hours did you sleep last night?

----

4. What time did you wake up today?

----

5. Did you have your breakfast? :

Yes

No

6. How many hours of lectures did you attend today?

----

7. When did you eat your last meal?

Next

Inventory Task

Instructions

In today's experiment, you will be making **inventory management decisions** for an enterprise called 'S-Store'. S-Store sells coffee makers. You will perform this role for a sequence of 6 years. Every month you will decide how many coffee makers to order from the coffee maker supplier. Your earnings in this experiment will be proportional to the total profitability of S-Store. S-store will sell a new coffee maker model every year. Thus in the first month of a year your inventory always starts from zero. Further, any coffee makers remaining in inventory at the end of month 12 will be disposed of. To summarise, you will be making 12 monthly decisions for a year, and you will do this for 6 years in total.

You will have up to **30 seconds** to complete your task for each month. Year 0 is a practice round, and you will have up to 20 seconds to complete the task for each month. You should use this as an opportunity to familiarize yourself with the software and decision tasks. If you don't finish within the time allowed, the computer will automatically execute the remaining month(s) sales with the existing inventory. You will not be able to add inventory. A 'wait page' displays automatically if you spend less than the allowed time in a year. You will only be able to proceed to the next year when the remaining time runs out.

Before the decision making portion of the experiment begins, there will be a **Test** consisting of 7 simple questions to check your understanding of the task. Please answer the questions carefully. If you missed 3 or more questions, you would be asked review the correct answers before you can proceed to the task.

Payment

Year 0 is a practice round, and you will receive no earnings from your decisions in this year. For

Years 1 through 5, your earnings will accumulate across years. At the end of the experiment you will be paid £5 show-up fee and your accumulated earnings, converted to Pounds. Note, negative profit may occur if poor coffee maker ordering decisions are made. To ensure that no one will leave the experiment with a payment less than £5, a negative total profit made in Year 1 to Year 5 will be treated as 0 earnings.

B.1.2 Background Information

[The following Background Information section shows up on every decision page.]

Your Role:

S-Store is open 360 days per year. You are the inventory manager for S-Store. In your role, you will control S-Store's inventory level which determines the store's total profits.

We now explain how S-Stores, and correspondingly you, earns profit. While we are explaining how the calculations are made, during the decision tasks the computer will carry out these calculations and report the results to you.

S-Store sells coffee makers at a price of **5 ECU** per unit. S-Store can sell up to **20 coffee makers per month**. A coffee maker can only be sold if there is a unit held in inventory. If you hold 20 or more units in inventory at the start of the month, S-Store will sell 20 coffee makers that month. However, if there are less than 20 units held in inventory at the start of the month then S-Store will only sell that amount. For example, if there are 2 units held in inventory at the beginning of a month then S-Store only sells 2 units that month. S-Store's sales revenue for a month is calculated as follows:

Sales revenue = 5 ECU * Number of units sold.

Your job is to manage the store's inventory levels by each month choosing an inventory order. Prior to the start of each month you can order coffee makers from the supplier to add to the inventory. Your inventory management determines the S-Store's total costs. S-Store pays two types of costs. One is the **ordering cost**. Every time you order a positive amount you have to pay an order cost. This ordering cost is **80 ECU**, and does not depend upon the size of the order. If you order zero coffee makers then you do not pay the 80 ECU ordering cost. Holding coffee makers in inventory is costly so S-Store pays a monthly **inventory holding cost**. S-Store pay's monthly inventory holding cost is based on the average number of coffee makers held in inventory multiplied by the per unit monthly inventory holding cost of **0.5 ECU**. This is calculated as follows:

Inventory holding costs = 0.5 ECU * (Opening inventory + Order Quantity + Closing inventory)/2.

Calculation of S-Store's profits

Profits = Sales revenue - Ordering costs - Inventory holding costs

Your monthly earnings are equal to S-Store's monthly profits.

Examples:

1. Alice's closing inventory of last month is 40 units, she placed an order of 0 units in this month. The demand for each month is 20 units.
   She made sales of 20 units.
   Her closing inventory of this month is $40 - 20 = 20$ units.
   Her **profit** in this month is equal to: $5 * 20 - 0 - 0.5 * (40 + 0 + 20)/2 = 85$.
2. Alice's closing inventory of last month is 10 units, she placed an order of 9 units in this month. The demand for each month is 20 units.

She only made sales of 19 units.
Her closing inventory of this month is 0 units.
Her **profit** in this month is equal to: $5 * 19 - 80 - 0.5 * (10 + 9 + 0)/2 = 10.25$.

B.1.3 Multiple Choice Questions prior to Inventory Task

There are a couple of questions for you before the task, please use the information below:

The demand for each month is 20 units.
Price of each coffee maker is 5.
Ordering cost is 80 per order.
Monthly inventory holding cost is 0.5 per unit.

Question 1 of 7
If the inventory level was 5 and you ordered 0 units. How many units will you SELL this month?

A 0
B 5
C 10
D 15

Question 2 of 7
If the inventory level was 0 and you ordered 25 units. How many units will you SELL this month?

A 0
B 10
C 20
D 25

Question 3 of 7
If you made sales of 20 units. What will be your SALES REVENUE this month?

A 0
B 20
C 80
D 100

Question 4 of 7
If you ordered 0 units. What will be your ORDERING COST this month?

A 0
B 0.5
C 80
D 100

Question 5 of 7
If you ordered 1 unit. What will be your ORDERING COST this month?

A 0
B 0.5
C 80
D 100

Question 6 of 7
If the inventory level was 0 and you ordered 20 units. You made sales of 20 units. What will be your HOLDING COST this month?

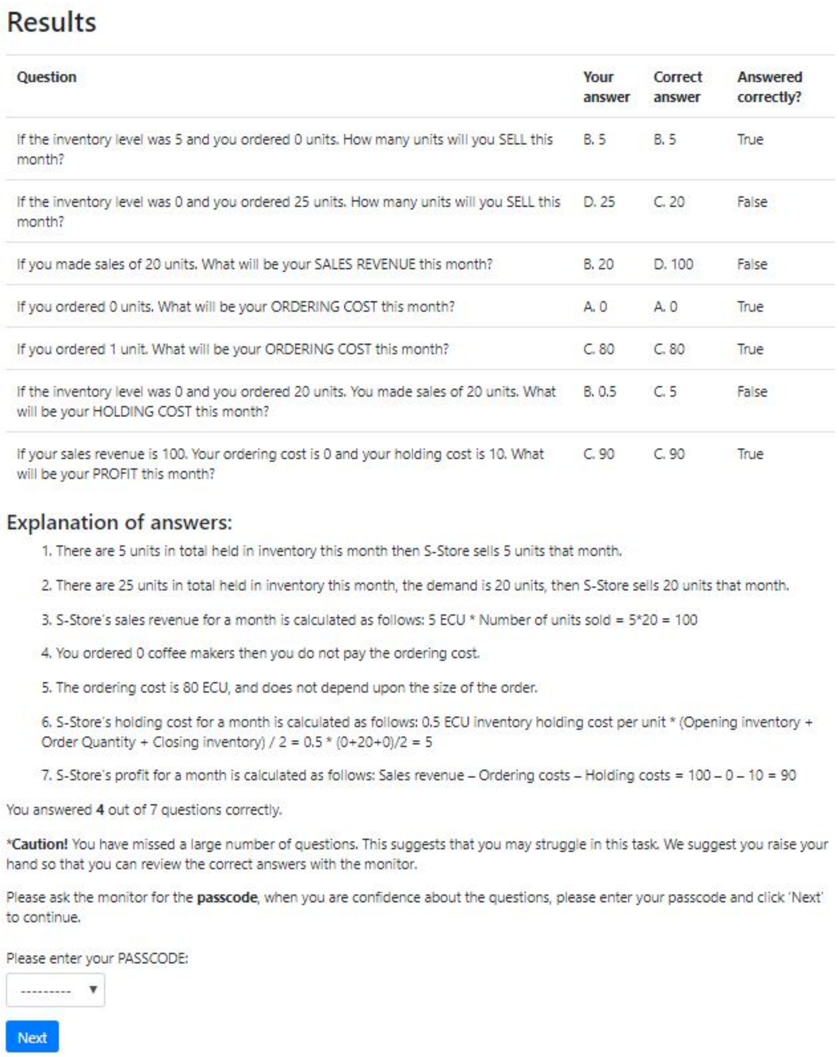
A 0
B 0.5
C 5
D 10

Question 7 of 7
If your sales revenue is 100. Your ordering cost is 0 and your holding cost is 10. What will be your PROFIT this month?

A 10
B 80
C 90
D 100

Figure 24 shows the result page of the multiple choice questions when participants had given more than 2 incorrect answers. Under such circumstances, they had to raise their hands to go through incorrectly answered questions with a monitor in order to obtain a passcode to proceed to the decision tasks.

Figure 24: Result page of the Multiple Choice Questions when more than 2 incorrect answers were provided



Results

Question	Your answer	Correct answer	Answered correctly?
If the inventory level was 5 and you ordered 0 units. How many units will you SELL this month?	B. 5	B. 5	True
If the inventory level was 0 and you ordered 25 units. How many units will you SELL this month?	D. 25	C. 20	False
If you made sales of 20 units. What will be your SALES REVENUE this month?	B. 20	D. 100	False
If you ordered 0 units. What will be your ORDERING COST this month?	A. 0	A. 0	True
If you ordered 1 unit. What will be your ORDERING COST this month?	C. 80	C. 80	True
If the inventory level was 0 and you ordered 20 units. You made sales of 20 units. What will be your HOLDING COST this month?	B. 0.5	C. 5	False
If your sales revenue is 100. Your ordering cost is 0 and your holding cost is 10. What will be your PROFIT this month?	C. 90	C. 90	True

Explanation of answers:

1. There are 5 units in total held in inventory this month then S-Store sells 5 units that month.
2. There are 25 units in total held in inventory this month, the demand is 20 units, then S-Store sells 20 units that month.
3. S-Store's sales revenue for a month is calculated as follows: 5 ECU * Number of units sold = 5*20 = 100
4. You ordered 0 coffee makers then you do not pay the ordering cost.
5. The ordering cost is 80 ECU, and does not depend upon the size of the order.
6. S-Store's holding cost for a month is calculated as follows: 0.5 ECU inventory holding cost per unit * (Opening inventory + Order Quantity + Closing inventory) / 2 = 0.5 * (0+20+0)/2 = 5
7. S-Store's profit for a month is calculated as follows: Sales revenue – Ordering costs – Holding costs = 100 – 0 – 10 = 90

You answered **4** out of 7 questions correctly.

***Caution!** You have missed a large number of questions. This suggests that you may struggle in this task. We suggest you raise your hand so that you can review the correct answers with the monitor.

Please ask the monitor for the **passcode**, when you are confidence about the questions, please enter your passcode and click 'Next' to continue.

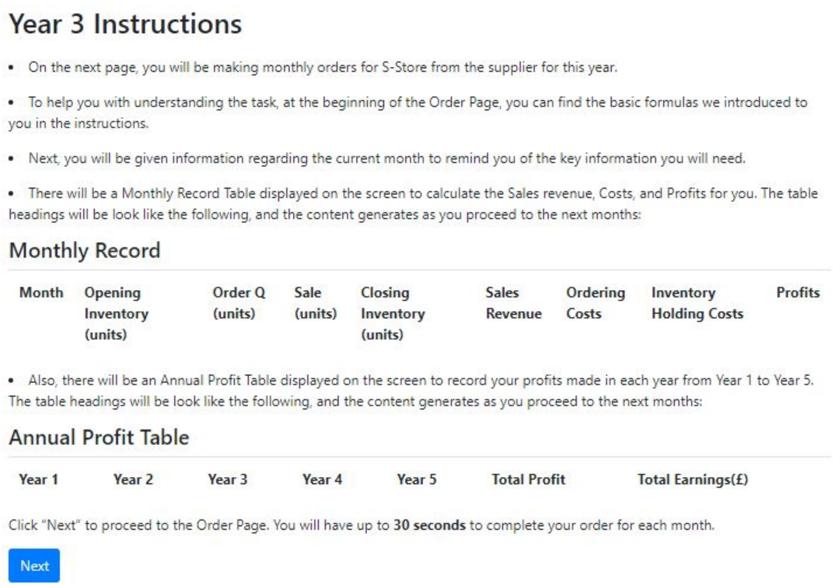
Please enter your PASSCODE:

Next

B.1.4 Inventory Task Interface

Prior to each year's inventory decision tasks, a mini-instruction page (see Figure 25 for an example) appears.

Figure 25: An example of the mini-instruction page prior to each year's inventory decision tasks



Year 3 Instructions

- On the next page, you will be making monthly orders for S-Store from the supplier for this year.
- To help you with understanding the task, at the beginning of the Order Page, you can find the basic formulas we introduced to you in the instructions.
- Next, you will be given information regarding the current month to remind you of the key information you will need.
- There will be a Monthly Record Table displayed on the screen to calculate the Sales revenue, Costs, and Profits for you. The table headings will be look like the following, and the content generates as you proceed to the next months:

Monthly Record

Month	Opening Inventory (units)	Order Q (units)	Sale (units)	Closing Inventory (units)	Sales Revenue	Ordering Costs	Inventory Holding Costs	Profits

- Also, there will be an Annual Profit Table displayed on the screen to record your profits made in each year from Year 1 to Year 5. The table headings will be look like the following, and the content generates as you proceed to the next months:

Annual Profit Table

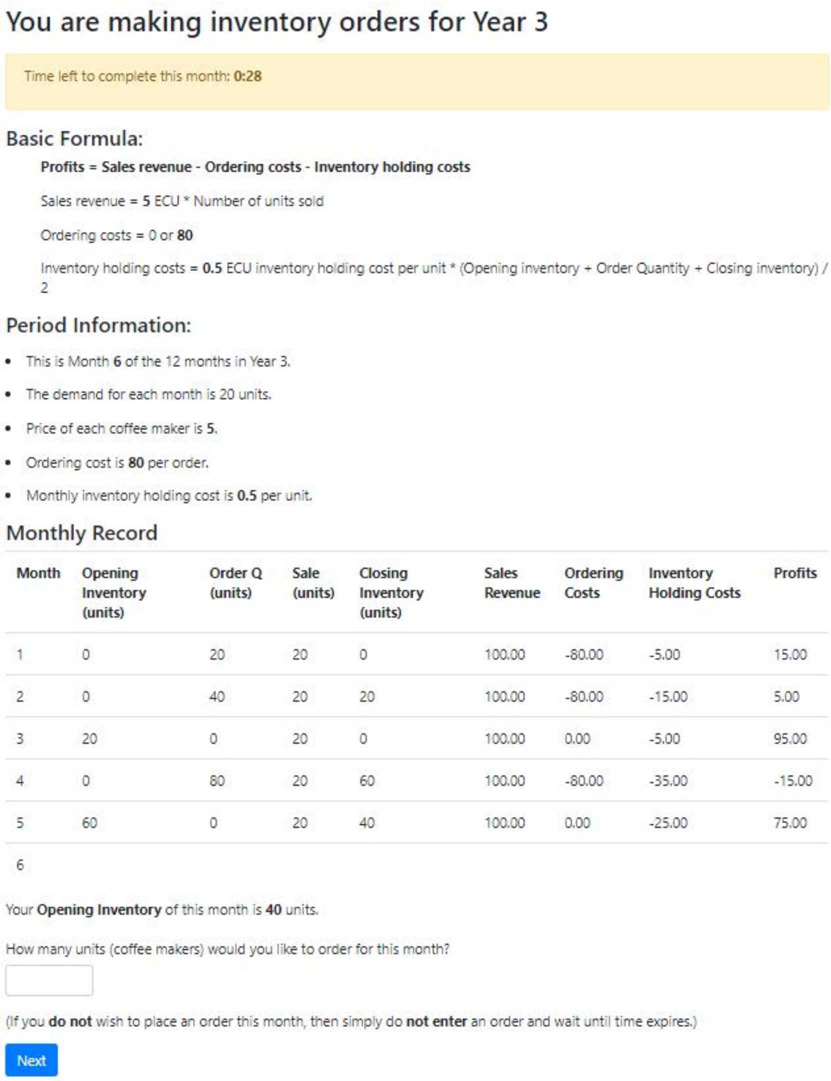
Year 1	Year 2	Year 3	Year 4	Year 5	Total Profit	Total Earnings(£)

Click "Next" to proceed to the Order Page. You will have up to **30 seconds** to complete your order for each month.

Next

An example of the ordering decision page is shown in Figure 26. Order quantities, costs, and profits of previous months are also displayed on the page. A participant needs to use the keyboard provided to enter his decision of order quantity for each month, if the decision is a positive order. In a case when a participant decides not to order for this month, he has to leave the box blank and waits on the decision page until time runs out, as number 0 is not allowed to be entered.

Figure 26: An example of inventory decision task page



You are making inventory orders for Year 3

Time left to complete this month: **0:28**

Basic Formula:

Profits = Sales revenue - Ordering costs - Inventory holding costs

Sales revenue = **5** ECU * Number of units sold

Ordering costs = 0 or **80**

Inventory holding costs = **0.5** ECU inventory holding cost per unit * (Opening inventory + Order Quantity + Closing inventory) / 2

Period Information:

- This is Month **6** of the 12 months in Year 3.
- The demand for each month is 20 units.
- Price of each coffee maker is **5**.
- Ordering cost is **80** per order.
- Monthly inventory holding cost is **0.5** per unit.

Monthly Record

Month	Opening Inventory (units)	Order Q (units)	Sale (units)	Closing Inventory (units)	Sales Revenue	Ordering Costs	Inventory Holding Costs	Profits
1	0	20	20	0	100.00	-80.00	-5.00	15.00
2	0	40	20	20	100.00	-80.00	-15.00	5.00
3	20	0	20	0	100.00	0.00	-5.00	95.00
4	0	80	20	60	100.00	-80.00	-35.00	-15.00
5	60	0	20	40	100.00	0.00	-25.00	75.00
6								

Your **Opening Inventory** of this month is **40** units.

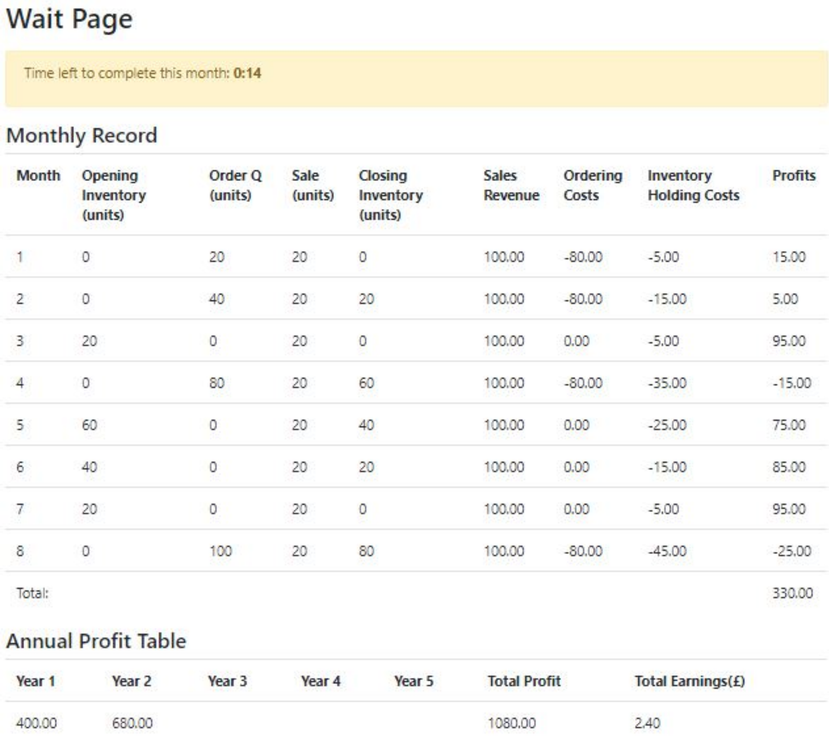
How many units (coffee makers) would you like to order for this month?

(If you **do not** wish to place an order this month, then simply do **not enter** an order and wait until time expires.)

Next

If a participant entered a positive order quantity and clicked "Next" before the timer had run out, he cannot proceed to next month decision page. A wait page (Figure 27) with information on previous months and previous years will appear instead.

Figure 27: An example of wait page when participants make monthly order decision before time runs out



Wait Page

Time left to complete this month: **0:14**

Monthly Record

Month	Opening Inventory (units)	Order Q (units)	Sale (units)	Closing Inventory (units)	Sales Revenue	Ordering Costs	Inventory Holding Costs	Profits
1	0	20	20	0	100.00	-80.00	-5.00	15.00
2	0	40	20	20	100.00	-80.00	-15.00	5.00
3	20	0	20	0	100.00	0.00	-5.00	95.00
4	0	80	20	60	100.00	-80.00	-35.00	-15.00
5	60	0	20	40	100.00	0.00	-25.00	75.00
6	40	0	20	20	100.00	0.00	-15.00	85.00
7	20	0	20	0	100.00	0.00	-5.00	95.00
8	0	100	20	80	100.00	-80.00	-45.00	-25.00
Total:								330.00

Annual Profit Table

Year 1	Year 2	Year 3	Year 4	Year 5	Total Profit	Total Earnings(£)
400.00	680.00				1080.00	2.40

After 12 months' decisions have been made, an end of the year result page which looks similar with the wait page in Figure 27 appears to provide an overview of their sales, revenue, costs and profits for every months, annual profits for the previous years, and accumulated earnings in pounds.

B.1.5 Post-Experimental Survey

Participants were asked to fill a simple questionnaire at the end of the experiment for us to collect some demographic information (Figure 28).

Figure 28: Post-Experimental Survey

Questionnaire

Please answer the following questions.

1. Have you had a course on supply chain management? :

- Yes
- No

2. What is your age?

----

3. What is your gender?

- Male
- Female

4. What is your country of citizenship?

----

5. Please indicate your current level of education:

- Undergraduate
- Postgraguate

6. Are you a native English speaker? :

- Yes
- No

Next

C Interview Transcripts

The motivation of the following interviews is to understand the disconnect of academic research and real-world scenarios. We intent to provide a wider application of our experimental research. In-depth conversations with inventory managers helped us to find better use case of the interventions that we have implemented in the experiments to improve participants performance, therefore, potentially improve the operations in the industry. As a result, we have proved the managerial value of our research and the importance of the EOQ model in the industry of non-durable goods.

The interviews were conducted during the period from April 2017 to Dec 2018 over video conference meetings. Our interviewees have agreed to publish the following part of our interview script which does not contain sensitive business data and information.

The following interview questions were answered by Mr Wu, the representative of inventory managers from an international furniture company based in Jiangsu, China.

1. Can you please briefly tell us about the inventory volume that you are managing?

Answer: We are a furniture chain brand, headquarters in Jiangsu, China. We have over 1000 retail stores across China that sells a large variety of products not limited to furniture, including stationaries and decorations. Our annual sales volume is around 30 million units. In terms of the demand, our demand keeps at a constant level on an annual basis, but there are peaks and off seasons every year.

2. What are the key responsibilities of an inventory manager?

Answer: We have several inventory managers who in charge of making the ordering decisions, the managers get the information of future demands from the marketing department. Inventory managers also have to count the inventory level every day, and update the inventory account records on a monthly basis.

3. What are the main sources of inventory managers' cognitive load in their daily job?

Answer: Generally speaking, inventory managers' cognitive stress come from two aspects, which are multitasking and workplace distractions. Multitasking: each of the inventory managers is managing multi-products, and the products can be either finished or semi-finished, paid in full or just the deposit. Distractions: the inventory managers will have to look after inventory coming in and out the warehouse, answering phones and communicate with other departments.

4. What are the programmes or apps you use in business for supply chain management?

Answer: The inventory management software we use is called the ERP system. Managers use the system to keep track of the inventories in stock, sales and ordering activities. Systems are good at providing real time solutions and yielding impressive profits improvement for companies when dealing with the same environment repeatedly. But managers do not entirely rely on it. They manually count the inventory levels every day, this is what other companies do as well. In the case where firms launch new product lines, inventory managers are also responsible for updating the inventory account inputs in the system every month accordingly. Our company hire inexperienced people, the typical training period is round two months, including the training on using the software. Managers are responsible for evaluating and making actual ordering decisions to overwrite the deterministic solution from the system.

The following interview questions were answered by Ms Su, the representative of inventory managers from three Chinese national firms operating in furniture, clothing and household hardware industries.

1. What are the elements of workplace stress of inventory managers, how is stress and demand of cognitive resources put on them?

Answer: The major external source of stress at work for inventory managers are the communications with the production and planning department. Managers need to participant in the inventory planning process, determine the optimal quantity as well as the time schedule. The job is cognitively demanding because such decisions influence the business's capacity to produce and meet the demand.

2. To what extend does technology provide an intervention that allows them to manage more?

Answer: The inventory managers use ERP system as the predominant tool to assist their management of inventories. They rely on the system's records to keep track of the inventories in stock, sales and ordering activities. However, they tend to coordinate with the manufacturers and make actual ordering decisions to overwrite the deterministic solution from the system. In the case of introducing new systems or new products launch, managers need to carry daily reviews on the system with the assistance of manual accounting. After the first year trail, managers stop doing manual accounting and rely entirely on the system. They switch from daily reviews to monthly reviews; account balancing within the predetermined error margin can be accepted.

3. Tell me about the type of people that you are look for to hire in this, what sort of scales are important, the rank of these abilities that are important?

Answer: The top three employment attributes that we are looking when hiring for inventory managers are adaptability and flexibility, reasoning and problem solving skills, and stress management skills. We also access their ability of making plans efficiently, whether they understand the concept of the ERP system, whether they are trustworthy professionals, and if they can improve their record-keeping and analytical skills in training. Workplace gender inequality is quite a common problem in supply chain management. Over 90% of the inventory managers in our companies are male managers, we are currently working to minimize the potential discrimination arising from current recruitment processes.